(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
15 March 2001 (15.03.2001)

PCT

(10) International Publication Number
WO 01/18050 A2

(51) International Patent Classification[7]: C07K 14/47, G01N 33/68, C12N 15/63, C07K 16/28

(21) International Application Number: PCT/US00/24821

(22) International Filing Date:
8 September 2000 (08.09.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/393,634 10 September 1999 (10.09.1999) US
09/510,332 22 February 2000 (22.02.2000) US

(71) Applicants *(for all designated States except US)*: **THE REGENTS OF THE UNIVERSITY OF CALIFORNIA** [US/US]; 1111 Franklin Street, Oakland, CA 94607 (US). **THE GOVERNMENT OF THE UNITED STATES OF AMERICA, as represented by THE SECRETARY, DEPARTMENT OF HEALTH AND HUMAN SERVICES** [US/US]; Suite 325, 6011 Executive Boulevard, Bethesda, MD 20852 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **ZUKER, Charles, S.** [US/US]; 4778 Thurston Place, San Diego, CA 92130 (US). **ADLER, Jon, Elliot** [US/US]; 1730 P. St. N.W., Washington, DC 20036 (US). **RYBA, Nick** [GB/US]; 9202 Lundigen Court, Bethesda, MD 20817 (US). **MUELLER, Ken** [US/US]; Apartment H1, 9585 Genesee Av., San Diego, CA 92121 (US). **HOON, Mark** [GB/US]; 4218 Warner Street, Kensington, MD 20895 (US).

(74) Agents: **FOLLETTE, Peter, J.** et al.; Townsend and Townsend and Crew LLP, 8th Floor, Two Embarcadero Center, San Francisco, CA 94111-3834 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *Without international search report and to be republished upon receipt of that report.*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/18050 A2

(54) Title: T2R, A NOVEL FAMILY OF TASTE RECEPTORS

(57) Abstract: The invention provides nucleic acid and amino acid sequences for a novel family of taste transduction G-protein coupled receptors, antibodies to such receptors, methods of detecting such nucleic acids and receptors, and methods of screening for modulators of taste transduction G-protein coupled receptors.

T2R, A NOVEL FAMILY OF TASTE RECEPTORS

CROSS-REFERENCES TO RELATED APPLICATIONS

This application claims priority to and is a continuation-in-part of USSN 09/393,634, filed September 10, 1999, which is herein incorporated by reference in its entirety.

STATEMENT AS TO FEDERALLY SPONSORED RESEARCH AND DEVELOPMENT

This invention was made with government support under Grant No. 5R01 DC03160, awarded by the National Institutes of Health. The government has certain rights in this invention.

FIELD OF THE INVENTION

The invention provides isolated nucleic acid and amino acid sequences of taste cell specific G-protein coupled receptors, antibodies to such receptors, methods of detecting such nucleic acids and receptors, and methods of screening for modulators of taste cell specific G-protein coupled receptors.

BACKGROUND OF THE INVENTION

Taste transduction is one of the most sophisticated forms of chemotransduction in animals (*see, e.g.*, Margolskee, *BioEssays* 15:645-650 (1993); Avenet & Lindemann, *J. Membrane Biol.* 112:1-8 (1989)). Gustatory signaling is found throughout the animal kingdom, from simple metazoans to the most complex of vertebrates; its main purpose is to provide a reliable signaling response to non-volatile ligands. Each of these modalities is though to be mediated by distinct signaling pathways mediated by receptors or channels, leading to receptor cell depolarization, generation of a receptor or action potential, and release of neurotransmitter at gustatory afferent neuron synapses (*see, e.g.*, Roper, *Ann. Rev. Neurosci.* 12:329-353 (1989)).

Mammals are believed to have five basic taste modalities: sweet, bitter, sour, salty, and umami (the taste of monosodium glutamate) (*see, e.g.*, Kawamura & Kare, *Introduction to Umami: A Basic Taste* (1987); Kinnamon & Cummings, *Ann. Rev. Physiol.* 54:715-731(1992); Lindemann, *Physiol. Rev.* 76:718-766 (1996); Stewart *et al.*, *Am. J. Physiol.* 272:1-26 (1997)). Extensive psychophysical studies in humans have reported that different regions of the tongue display different gustatory preferences (*see, e.g.*, Hoffmann, *Menchen. Arch. Path. Anat. Physiol.* 62:516-530 (1875); Bradley *et al.*, *Anatomical Record* 212: 246-249 (1985); Miller & Reedy, *Physiol. Behav.* 47:1213-1219 (1990)). Also, numerous physiological studies in animals have shown that taste receptor cells may selectively respond to different tastants (*see, e.g.*, Akabas *et al.*, *Science* 242:1047-1050 (1988); Gilbertson *et al.*, *J. Gen. Physiol.* 100:803-24 (1992); Bernhardt *et al.*, *J. Physiol.* 490:325-336 (1996); Cummings *et al.*, *J. Neurophysiol.* 75:1256-1263 (1996)).

In mammals, taste receptor cells are assembled into taste buds that are distributed into different papillae in the tongue epithelium. Circumvallate papillae, found at the very back of the tongue, contain hundreds (mice) to thousands (human) of taste buds and are particularly sensitive to bitter substances. Foliate papillae, localized to the posterior lateral edge of the tongue, contain dozens to hundreds of taste buds and are particularly sensitive to sour and bitter substances. Fungiform papillae containing a single or a few taste buds are at the front of the tongue and are thought to mediate much of the sweet taste modality.

Each taste bud, depending on the species, contains 50-150 cells, including precursor cells, support cells, and taste receptor cells (*see, e.g.*, Lindemann, *Physiol. Rev.* 76:718-766 (1996)). Receptor cells are innervated at their base by afferent nerve endings that transmit information to the taste centers of the cortex through synapses in the brain stem and thalamus. Elucidating the mechanisms of taste cell signaling and information processing is critical for understanding the function, regulation, and "perception" of the sense of taste.

Although much is known about the psychophysics and physiology of taste cell function, very little is known about the molecules and pathways that mediate these sensory signaling responses (reviewed by Gilbertson, *Current Opin. Neurobiol.* 3:532-539 (1993)). Electrophysiological studies suggest that sour and salty tastants modulate taste cell function by direct entry of H^+ and Na^+ ions through specialized membrane channels on the apical surface of the cell. In the case of sour compounds, taste cell

depolarization is hypothesized to result from H^+ blockage of K^+ channels (*see, e.g.*, Kinnamon *et al.*, *Proc. Nat'l Acad. Sci. USA* 85: 7023-7027 (1988)) or activation of pH-sensitive channels (*see, e.g.*, Gilbertson *et al.*, *J. Gen. Physiol.* 100:803-24 (1992)); salt transduction may be partly mediated by the entry of Na^+ via amiloride-sensitive Na^+ channels (*see, e.g.*, Heck *et al.*, *Science* 223:403-405 (1984); Brand *et al.*, *Brain Res.* 207-214 (1985); Avenet *et al.*, *Nature* 331: 351-354 (1988)).

Sweet, bitter, and umami transduction are believed to be mediated by G-protein-coupled receptor (GPCR) signaling pathways (*see, e.g.*, Striem *et al.*, *Biochem. J.* 260:121-126 (1989); Chaudhari *et al.*, *J. Neuros.* 16:3817-3826 (1996); Wong *et al.*, *Nature* 381: 796-800 (1996)). Confusingly, there are almost as many models of signaling pathways for sweet and bitter transduction as there are effector enzymes for GPCR cascades (*e.g.*, G protein subunits, cGMP phosphodiesterase, phospholipase C, adenylate cyclase; *see, e.g.*, Kinnamon & Margolskee, *Curr. Opin. Neurobiol.* 6:506-513 (1996)). However, little is known about the specific membrane receptors involved in taste transduction, or many of the individual intracellular signaling molecules activated by the individual taste transduction pathways. Identification of such molecules is important given the numerous pharmacological and food industry applications for bitter antagonists, sweet agonists, and other modulators of taste.

One taste-cell specific G protein that has been identified is called Gustducin (McLaughin *et al.*, *Nature* 357:563-569 (1992)). This protein is proposed to be involved in the detection of certain bitter and sweet tastes since gustducin knockout mice show decreased sensitivity to some sweet and bitter tastants (Wong *et al.*, *Nature* 381:796-800 (1996)), and because gustducin is expressed in a significant subset of cells from all types of taste papillae (McLaughin *et al.*, *Nature* 357:563-569 (1992)). In addition, gustducin can be activated *in vitro* by stimulating taste membranes with bitter compounds, likely through the activation of bitter receptors (Ming *et al, PNAS* 95:8933-8938 (1998)).

Recently, two novel GPCRs were identified and found to be specifically expressed in taste cells. While these receptor proteins, called TR1 and TR2, appear to be directly involved in taste reception (Hoon *et al.*, *Cell* 96:541-551 (1999)), they are only expressed in a fraction of mammalian taste receptor cells. For example, neither of the genes are extensively expressed in Gustducin-expressing cells. Thus, it is clear that additional taste-involved GPCRs remain to be discovered.

Gene studies in mammals have identified numer loci that are involved in the detection of taste. For example, psychophysical tasting studies have shown that humans can be categorized as tasters, non-tasters, and super-tasters for the bitter substance PROP (6-n-propylthiouracil), and that PROP tasting may be conferred by a dominant allele, with non-tasters having two recessive alleles and tasters having at least one dominant allele (*see* Bartoshuk *et al., Physiol Behav* 56(6):1165-71; 58:203-204 (1994)). Recently, a locus involved in PROP tasting has been mapped to human interval 5p15 (Reed *et al., Am. J. Hum. Genet.*, 64(5):1478-80 (1999)). The PROP tasting gene present at the 5p15 locus has yet to be described, however.

In addition, a number of genes involved in taste have been mapped in mice. For example, a cluster of genes involved in bitter-taste detection has been mapped to a region of chromosome 6 in mice (Lush *et al., Genet Res.* 66:167-174 (1995)).

The identification and isolation of novel taste receptors and taste signaling molecules would allow for new methods of pharmacological and genetic modulation of taste transduction pathways. For example, the availability of receptor and channel molecules would permit the screening for high affinity agonists, antagonists, inverse agonists, and modulators of taste cell activity. Such taste modulating compounds would be useful in the pharmaceutical and food industries to customize taste. In addition, such taste cell specific molecules can serve as invaluable tools in the generation of taste topographic maps that elucidate the relationship between the taste cells of the tongue and taste sensory neurons leading to taste centers in the brain.

SUMMARY OF THE INVENTION

The present invention thus provides novel nucleic acids encoding a family of taste-cell specific G-protein coupled receptors. These nucleic acids and the polypeptides that they encode are referred to as the "T2R" family of G-protein coupled taste receptors. These receptors are also referred to as the "SF" family of G-protein coupled taste receptors. This novel family of GPCRs includes components of the taste transduction pathway. In particular, members of this family are involved in the detection of bitter tastes.

In one aspect, the present invention provides a method for identifying a compound that modulates taste signaling in taste cells, the method comprising the steps of: (i) contacting a taste transduction G-protein coupled receptor polypeptide with the compound, the polypeptide comprising at least about 50% amino acid identity to a

sequence selected from the group consisting of SEQ ID NO:166, SEQ ID NO:167, SEQ ID NO:168, SEQ ID NO:169, SEQ ID NO:170, and SEQ ID NO:171; and (ii) determining the functional effect of the compound upon the polypeptide.

In another aspect, the present invention provides a method for identifying a compound that modulates taste signaling in taste cells, the method comprising the steps of: (i) contacting a taste transduction G-protein coupled receptor polypeptide with the compound, the polypeptide comprising greater than about 60% amino acid sequence identity to a sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5; SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:35, SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:44, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:62, SEQ ID NO:64, SEQ ID NO:65, SEQ ID NO:66, SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, SEQ ID NO:76, SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81, SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID NO:158, SEQ ID NO:160, SEQ ID NO:162, and SEQ ID NO:164; and (ii) determining the functional effect of the compound upon the polypeptide.

In another aspect, the present invention provides a method for identifying a compound that modulates taste signaling in taste cells, the method comprising the steps of: (i) contacting a polypeptide comprising an extracellular domain or transmembrane region, or combination thereof, of a taste transduction G-protein coupled receptor with the compound, the extracellular domain or transmembrane region comprising greater than about 60% amino acid sequence identity to the extracellular domain or transmembrane

region of a polypeptide comprising a sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5; SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:35, SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:44, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:62, SEQ ID NO:64, SEQ ID NO:65, SEQ ID NO:66, SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, SEQ ID NO:76, SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81, SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID NO:158, SEQ ID NO:160, SEQ ID NO:162, and SEQ ID NO:164; and (ii) determining the functional effect of the compound upon the extracellular domain or transmembrane region.

In one embodiment, the polypeptide has G-protein coupled receptor activity. In another embodiment, the functional effect is a chemical effect. In another embodiment, the functional effect is a physical effect. In another embodiment, the functional effect is determined by measuring binding of the compound to an extracellular domain of the polypeptide. In another embodiment, the functional effect is determined by measuring radiolabeled GTP binding to the polypeptide. In another embodiment, the polypeptide is recombinant. In another embodiment, the polypeptide comprises an extracellular domain or transmembrane region or a combination of an extracellular domain and transmembrane region that is covalently linked to a heterologous polypeptide, forming a chimeric polypeptide. In another embodiment, the polypeptide is linked to a solid phase, either covalently or non-covalently. In another embodiment, the polypeptide is from a rat, a mouse, or a human.

In another embodiment, the polypeptide is expressed in a cell or a cell membrane. In another embodiment, the cell is a eukaryotic cell. In another embodiment, the functional effect is measured by determining changes in the electrical activity of a cell expressing the polypeptide. In another embodiment, the functional effect of the compound upon the polypeptide is determined by measuring changes in intracellular cAMP, cGMP, IP3, or Ca^{2+} in a cell expressing the polypeptide. In another embodiment, a change in intracellular Ca^{2+} in the cell is detected by detecting FURA-2 dependent fluorescence in the cell. In another embodiment, the cell is a eukaryotic cell. In another embodiment, the cell is an HEK-293 cell. In another embodiment, the polypeptide is a fusion protein comprising at least about 20 consecutive N-terminal amino acids of a rhodopsin protein. In another embodiment, the rhodopsin protein is a bovine rhodopsin. In another embodiment, the cell comprises Gα15. In another embodiment, the polypeptide is expressed in a cell, and the polypeptide is contacted with the compound in the presence of a bitter tastant, wherein a difference in the functional effect of the bitter tastant on the cell in the presence of the compound and the functional effect of the bitter tastant on the cell in the absence of the compound indicates that the compound is capable of modulating taste signaling in taste cells.

In another embodiment, the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5; SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:35, SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:44, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:62, SEQ ID NO:64, SEQ ID NO:65, SEQ ID NO:66, SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, SEQ ID NO:76, SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81, SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID

NO:123, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID NO:158, SEQ ID NO:160, SEQ ID NO:162, and SEQ ID NO:164.

In another aspect, the present invention provides an isolated nucleic acid encoding a taste transduction G-protein coupled receptor, the receptor comprising greater than about 50% amino acid sequence identity to a sequence selected from the group consisting of SEQ ID NO:166, SEQ ID NO:167, SEQ ID NO:168, SEQ ID NO:169, SEQ ID NO:170, and SEQ ID NO:171.

In another aspect, the present invention provides an isolated nucleic acid encoding a taste transduction G-protein coupled receptor, wherein the nucleic acid is amplified by primers that selectively hybridize to the same sequence as degenerate primer sets encoding amino acid sequences selected from the group consisting of SEQ ID NO:166, SEQ ID NO:167, SEQ ID NO:168, SEQ ID NO:169, SEQ ID NO:170, and SEQ ID NO:171.

In another aspect, the present invention provides an isolated nucleic acid encoding a taste transduction G-protein coupled receptor, the receptor comprising greater than about 60% amino acid sequence identity to a sequence selected from the group consisting of SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81, SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID NO:158, SEQ ID NO:160, SEQ ID NO:162, and SEQ ID NO:164.

In another aspect, the present invention provides an isolated nucleic acid encoding a taste transduction G-protein coupled receptor, wherein the nucleic acid specifically hybridizes under highly stringent conditions to a nucleic acid having a nucleotide sequence selected from the group consisting of SEQ ID NO:78, SEQ ID NO:80, SEQ ID NO:82, SEQ ID NO:84, SEQ ID NO:86; SEQ ID NO:88, SEQ ID NO:90, SEQ ID NO:92, SEQ ID NO:94, SEQ ID NO:96, SEQ ID NO:98, SEQ ID

NO:100, SEQ ID NO:102, SEQ ID NO:104 SEQ ID NO:106, SEQ ID NO:108, SEQ ID NO:110, SEQ ID NO:112, SEQ ID NO:114, SEQ ID NO:116, SEQ ID NO:118, SEQ ID NO:120, SEQ ID NO:120, SEQ ID NO:122, SEQ ID NO:124, SEQ ID NO:126, SEQ ID NO:128, SEQ ID NO:130, SEQ ID NO:132, SEQ ID NO:134, SEQ ID NO:136, SEQ ID NO:138, SEQ ID NO:140, SEQ ID NO:142, SEQ ID NO:144, SEQ ID NO:146, SEQ ID NO:148, SEQ ID NO:150, SEQ ID NO:152, SEQ ID NO:154, SEQ ID NO:156, SEQ ID NO:157, SEQ ID NO:159, SEQ ID NO:161, SEQ ID NO:163, and SEQ ID NO:165, but not to a nucleic acid having a nucleotide sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:20, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:34, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:41, SEQ ID NO:43, SEQ ID NO:45, SEQ ID NO:52, SEQ ID NO:54, SEQ ID NO:57, SEQ ID NO:61, and SEQ ID NO:63.

In another aspect, the present invention provides an isolated nucleic acid encoding a taste transduction G-protein coupled receptor, the receptor comprising greater than about 60% amino acid identity to a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81, SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID NO:158, SEQ ID NO:160, SEQ ID NO:162, and SEQ ID NO:164, wherein the nucleic acid selectively hybridizes under moderately stringent hybridization conditions to a nucleotide sequence having a nucleotide sequence selected from the group consisting of SEQ ID NO:78, SEQ ID NO:80, SEQ ID NO:82, SEQ ID NO:84, SEQ ID NO:86; SEQ ID NO:88, SEQ ID NO:90, SEQ ID NO:92, SEQ ID NO:94, SEQ ID NO:96, SEQ ID NO:98, SEQ ID NO:100, SEQ ID NO:102, SEQ ID NO:104 SEQ ID NO:106, SEQ ID NO:108, SEQ ID NO:110, SEQ ID NO:112, SEQ ID NO:114, SEQ ID NO:116, SEQ ID NO:118, SEQ ID NO:120, SEQ ID NO:120, SEQ ID NO:122, SEQ ID NO:124, SEQ ID NO:126, SEQ ID NO:128, SEQ ID NO:130, SEQ ID

NO:132, SEQ ID NO:134, SEQ ID NO:136, SEQ ID NO:138, SEQ ID NO:140, SEQ ID NO:142, SEQ ID NO:144, SEQ ID NO:146, SEQ ID NO:148, SEQ ID NO:150, SEQ ID NO:152, SEQ ID NO:154, SEQ ID NO:156, SEQ ID NO:157, SEQ ID NO:159, SEQ ID NO:161, SEQ ID NO:163, and SEQ ID NO:165 but not to a nucleic acid having a nucleotide sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:20, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:34, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:41, SEQ ID NO:43, SEQ ID NO:45, SEQ ID NO:52, SEQ ID NO:54, SEQ ID NO:57, SEQ ID NO:61, and SEQ ID NO:63.

In another aspect, the present invention provides an isolated nucleic acid encoding an extracellular domain or transmembrane region or a combination thereof of a taste transduction G-protein coupled receptor, the extracellular domain or transmembrane region having greater than about 60% amino acid sequence identity to the extracellular domain or transmembrane region of a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81, SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID NO:158, SEQ ID NO:160, SEQ ID NO:162, and SEQ ID NO:164.

In one embodiment, the nucleic acid encodes a receptor that specifically binds to polyclonal antibodies generated against a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81, SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID

NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID NO:158, SEQ ID NO:160, SEQ ID NO:162, and SEQ ID NO:164, but not to polyclonal antibodies generated against a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5; SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:35, SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:44, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:62, SEQ ID NO:64, SEQ ID NO:65, SEQ ID NO:66, SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, and SEQ ID NO:76.

In another embodiment, the nucleic acid encodes a receptor comprising an amino acid sequence selected from the group consisting of SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81, SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID NO:158, SEQ ID NO:160, SEQ ID NO:162, SEQ ID NO:164, SEQ ID NO:166, SEQ ID NO:167, SEQ ID NO:168, SEQ ID NO:169, SEQ ID NO:170, and SEQ ID NO:171.

In another embodiment, the nucleic acid comprises a nucleotide sequence selected from the group consisting of SEQ ID NO:78, SEQ ID NO:80, SEQ ID NO:82, SEQ ID NO:84, SEQ ID NO:86; SEQ ID NO:88, SEQ ID NO:90, SEQ ID NO:92, SEQ ID NO:94, SEQ ID NO:96, SEQ ID NO:98, SEQ ID NO:100, SEQ ID NO:102, SEQ ID NO:104 SEQ ID NO:106, SEQ ID NO:108, SEQ ID NO:110, SEQ ID NO:112, SEQ ID NO:114, SEQ ID NO:116, SEQ ID NO:118, SEQ ID NO:120, SEQ ID NO:120, SEQ ID NO:122, SEQ ID NO:124, SEQ ID NO:126, SEQ ID NO:128, SEQ ID NO:130, SEQ ID NO:132, SEQ ID NO:134, SEQ ID NO:136, SEQ ID NO:138, SEQ ID NO:140, SEQ ID

NO:142, SEQ ID NO:144, SEQ ID NO:146, SEQ ID NO:148, SEQ ID NO:150, SEQ ID NO:152, SEQ ID NO:154, SEQ ID NO:156, SEQ ID NO:157, SEQ ID NO:159, SEQ ID NO:161, SEQ ID NO:163, and SEQ ID NO:165.

In another embodiment, the nucleic acid encodes a receptor that has G-protein coupled receptor activity. In another embodiment, the nucleic acid is from a rat or a mouse.

In another embodiment, the nucleic acid encodes an extracellular domain or transmembrane region or combination thereof linked to a heterologous polypeptide, forming a chimeric polypeptide. In another embodiment, the nucleic acid encodes the extracellular domain of a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81, SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID NO:158, SEQ ID NO:160, SEQ ID NO:162, SEQ ID NO:164, SEQ ID NO:166, SEQ ID NO:167, SEQ ID NO:168, SEQ ID NO:169, SEQ ID NO:170, and SEQ ID NO:171.

In another aspect, the present invention provides an expression vector comprising any of the above nucleic acids. In another aspect, the present invention provides isolated cells comprising the expression vector.

In another aspect, the present invention provides an isolated taste transduction G-protein coupled receptor, the receptor comprising greater than about 50% amino acid sequence identity to a sequence selected from the group consisting of SEQ ID NO:166, SEQ ID NO:167, SEQ ID NO:168, SEQ ID NO:169, SEQ ID NO:170, and SEQ ID NO:171.

In another aspect, the present invention provides an isolated taste transduction G-protein coupled receptor, the receptor comprising greater than about 60% amino acid sequence identity to a sequence selected from the group consisting of SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81, SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID

NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID NO:158, SEQ ID NO:160, SEQ ID NO:162, and SEQ ID NO:164.

In one embodiment, the receptor comprises an amino acid sequence selected from the group consisting of SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81, SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID NO:158, SEQ ID NO:160, SEQ ID NO:162, SEQ ID NO:164, SEQ ID NO:166, SEQ ID NO:167, SEQ ID NO:168, SEQ ID NO:169, SEQ ID NO:170, and SEQ ID NO:171.

In another embodiment, the receptor specifically binds to polyclonal antibodies generated against a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81, SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID NO:158, SEQ ID NO:160, SEQ ID NO:162, and SEQ ID NO:164, but not to polyclonal antibodies generated against a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5; SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:24, SEQ

ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:35, SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:44, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:62, SEQ ID NO:64, SEQ ID NO:65, SEQ ID NO:66, SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, and SEQ ID NO:76. In another embodiment, the receptor has G-protein coupled receptor activity. In another embodiment, the receptor is from a rat or a mouse.

In another aspect, the present invention provides an isolated polypeptide comprising an extracellular domain or a transmembrane region or a combination thereof of a taste transduction G-protein coupled receptor, the extracellular domain or transmembrane region comprising greater than about 60% amino acid sequence identity to the extracellular domain or transmembrane region of a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81, SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID NO:158, SEQ ID NO:160, SEQ ID NO:162, and SEQ ID NO:164.

In one embodiment, the polypeptide encodes the extracellular domain or transmembrane region of a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81, SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID

NO:155, SEQ ID NO:158, SEQ ID NO:160, SEQ ID NO:162, SEQ ID NO:164, SEQ ID NO:166, SEQ ID NO:167, SEQ ID NO:168, SEQ ID NO:169, SEQ ID NO:170, and SEQ ID NO:171. In another embodiment, the extracellular domain or transmembrane region is covalently linked to a heterologous polypeptide, forming a chimeric polypeptide.

In one aspect, the present invention provides an antibody that selectively binds to the receptor comprising greater than about 60% amino acid sequence identity to a sequence selected from the group consisting of SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81, SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID NO:158, SEQ ID NO:160, SEQ ID NO:162, and SEQ ID NO:164.

In another aspect, the present invention provides an expression vector comprising a nucleic acid encoding a taste transduction G-protein coupled receptor, wherein the receptor is expressed in a taste cell, the receptor comprising greater than about 60% amino acid sequence identity to a sequence selected from the group consisting of SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81, SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID NO:158, SEQ ID NO:160, SEQ ID NO:162, and SEQ ID NO:164.

In another aspect, the present invention provides a host cell transfected with the expression vector.

In another aspect, the present invention provides an expression cassette comprising a polynucleotide sequence that encodes a human taste transduction G protein coupled receptor, operably linked to a heterologous promoter, wherein the receptor

comprises an amino acid sequence comprising greater than about [redacted]% amino acid sequence identity to a sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5; SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:35, SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:44, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:62, SEQ ID NO:64, SEQ ID NO:65, SEQ ID NO:66, SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, and SEQ ID NO:76.

In one embodiment, the receptor comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5; SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:35, SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:44, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:62, SEQ ID NO:64, SEQ ID NO:65, SEQ ID NO:66, SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, and SEQ ID NO:76.

In another aspect, the present invention provides an isolated eukaryotic cell comprising the expression cassette.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 demonstrates that Gα15 couples the activation of μ opioid receptor and mGluR1 receptor to the release of intracellular calcium. HEK-293 cells were transiently transfected with the Gαi coupled μ opioid receptor or the Gαq coupled mGluR1 receptor. Transfected cells containing Gα15 were assayed for increases in [Ca2+]i before (a, b) and after (c, d) the addition of receptor agonists: (c) 10μM DAMGO

and (d) 20 μM tran(±) 1-amino-1,3 cyclopentane dicarboxylic a(ACPD). Ligand- and receptor-dependent increase in [Ca2+]i were dependent on Gα15 (panels e, f). Scales indicate [Ca2+]i (nM) determined from FURA-2 emission ratios.

Figure 2 shows that the first 39 amino acids of bovine rhodopsin effectively targets T2Rs to the plasma membrane of HEK-293 cells. Immunofluorescence staining of non-permeabilized cells transfected with representative rho-T2R fusions was detected using an anti-rhodopsin mAb B6-30.

Figure 3 demonstrates that T2R receptors are stimulated by bitter compounds. HEK-293 cells were transfected with rho-mT2R5 (a, d, g), rho-hT2R4 (b, e, h), and rho-mT2R8 (c, f, i). Cells expressing mT2R5 were stimulated using 1.5 μM cycloheximide (d, g) and those expressing hT2R4 and mT2R8 with 1.5 mM denatonium (e, f, h, i). No increase in [Ca2+]i was observed in the absence of Gα15 (g - i); in contrast robust Gα15 dependent responses were observed in the presence of tastants (d - f); scales indicate [Ca2+]i (nM) determined from FURA-2 emission ratios. Line traces (j - l) show the kinetics of the [Ca2+]i changes for representative cells from panels (d - f); arrows indicate addition of tastants.

Figure 4 shows that mT2R5 is a taste receptor for cycloheximide. (a) HEK-293 cells expressing Gα15 and rho-mT2R5 were challenged with multiple pulses of 2 μM cycloheximide (CYX), 3 mM 6-n-propyl thiouracil (PROP) or 5 mM denatonium (DEN); dots and horizontal bars above the traces indicate the time and duration of tastant pulses. Cycloheximide triggers robust receptor activation. This experiment also illustrates desensitization to repeated stimulation or during sustained application of the stimulus. (b) Responses to cycloheximide are highly specific and are not observed after addition of buffer (CON) or high concentrations of other tastants. Abbreviations and concentrations used are: cycloheximide, CYX (5 μM); atropine, ATR (5 mM); brucine, BRU (5 mM); caffeic acid, CAFF (2 mM); denatonium, DEN (5 mM); epicatechin, (-)EPI (3 mM); phenyl thiocarbamide, PTC (3 mM); 6-n-propyl thiouracil, PROP (10 mM); saccharin, SAC (10 mM); strychnine, STR (5 mM); sucrose octaacetate, SOA (3 mM). Columns represent the mean ± s.e of at least six independent experiments. (c) The mT2R5 gene from taster (DBA/2-allele) and non-taster (C57BL/6-allele) strains mediate differential [Ca2+]i changes to pulses of cycloheximide. Horizontal bars depict the time and duration of the stimulus. 200 s was allowed to elapse between stimuli to ensure that cells were not desensitized due to the successive application of cycloheximide. (d)

Cycloheximide dose response of mT2R5. Changes in [Ca2+]i are shown as FURA-2 (F340/F380) ratios normalized to the response at 30 μM cycloheximide; points represent the mean ± s.e. of at least six determinations. The non-taster allele shows a marked decrease in cycloheximide sensitivity relative to the taster allele (EC50s of ~2.3 μM versus 0.5 μM, respectively).

Figure 5 shows that hT2R4 and mT2R8 respond to denatonium. HEK-293 cells expressing Gα15 were transiently transfected with hT2R4 or mT2R8 receptors and [Ca2+]i was monitored as shown in Figure 3. (a) An increase in [Ca2+]i could be induced by stimulation with denatonium but not by various other bitter compounds. Response profiles of (b) hT2R4 and (c) mT2R8 to a set of nine out of 55 different bitter and sweet tastants (see Experimental Procedures) are shown. CON refers to control buffer addition, NAR to 2mM naringin and LYS to 5mM lysine. Other abbreviations and concentrations are as reported in Figure 4. The mean FURA-2 fluorescence ratio (F340/F380) before and after ligand addition was obtained from 100 equal sized areas that included all responding cells. The values represent the mean± s.e. of at least 6 experiments.

Figure 6 demonstrates that cycloheximide taster and non-taster strains express different alleles of mT2R5. (a) Predicted transmembrane topology of mT2R5; amino-acid substitutions in the allele from non-taster strains are highlighted in red. The presence of only two alleles at this locus is not unexpected because the strains that share the same polymorphisms were derived from a common founder (Beck *et al.*, *Nat Genet* 24:23-55 (2000)). *In situ* hybridization showing expression of mT2R5 in subsets of cells in the circumvallate papilla of (b) a cycloheximide taster strain (DBA/2) and (c) a non-taster strain (C57BL/6); no strain specific differences in expression pattern were detected in taste buds from other regions of the oral cavity.

Figure 7 shows that mT2R5 activates gustducin in response to cycloheximide. (a) Insect larval cell membranes containing mT2R5 activate gustducin in the presence 300 μM cycloheximide but not without ligand (control) or in the presence of 1 mM atropine, brucine, caffeine, denatonium, phenylthiocarbamide, 6-n-propyl thiouracil, quinine, saccharin, strychnine, sucrose octaacetate. (b) Cycloheximide concentration dependence of gustducin activation by mT2R5 was fitted by single-site binding (Kd=14.8 + 0.9 μM).

Figure 8 provides a table including nucleic acid and protein sequences for a number of human, rat, and mouse T2R family members.

DETAILED DESCRIPTION OF THE INVENTION

I. Introduction

The present invention provides nucleic acids encoding a novel family of taste cell specific G-protein coupled receptors. These nucleic acids and the receptors that they encode are referred to as members of the "T2R" family of taste cell specific G protein coupled receptors. These taste cell specific GPCRs are components of the taste transduction pathway, *e.g.*, the bitter taste transduction pathway, and are involved in the taste detection of substances such as the bitter substances 6-n-propylthiouracil (PROP), sucrose octaacetate (soa), raffinose undecaacetate (roa), cycloheximide (cyx), denatonium, copper glycinate (Glb), and quinine (qui).

These nucleic acids provide valuable probes for the identification of taste cells, as the nucleic acids are specifically expressed in taste cells. For example, probes for T2R polypeptides and proteins can be used to identity taste cells present in foliate, circumvallate, and fungiform papillae, as well as taste cells present in the geschmackstreifen and epiglottis. In particular, T2R probes are useful to indentify bitter sensing, gustducin expressing taste cells. They also serve as tools for the generation of taste topographic maps that elucidate the relationship between the taste cells of the tongue and taste sensory neurons leading to taste centers in the brain. Furthermore, the nucleic acids and the proteins they encode can be used as probes to dissect taste-induced behaviors.

The invention also provides methods of screening for modulators, *e.g.*, activators, inhibitors, stimulators, enhancers, agonists, and antagonists, of these novel taste cell GPCRs. Such modulators of taste transduction are useful for pharmacological and genetic modulation of taste signaling pathways. These methods of screening can be used to identify high affinity agonists and antagonists of taste cell activity. These modulatory compounds can then be used in the food and pharmaceutical industries to customize taste, for example, to decrease the bitter taste of foods or drugs. Thus, the invention provides assays for taste modulation, where members of the T2R family act as direct or indirect reporter molecules for the effect of modulators on taste transduction. GPCRs can be used in assays, *e.g.*, to measure changes in ligand binding, ion concentration, membrane potential, current flow, ion flux, transcription, signal

transduction, receptor-ligand interactions, second messenger concentrations, *in vitro*, *in vivo*, and *ex vivo*. In one embodiment, members of the T2R family can be used as indirect reporters via attachment to a second reporter molecule such as green fluorescent protein (*see, e.g.*, Mistili & Spector, *Nature Biotechnology* 15:961-964 (1997)). In another embodiment, T2R family members are recombinantly expressed in cells, and modulation of taste transduction via GPCR activity is assayed by measuring changes in Ca^{2+} levels and other intracellular messages such as cAMP, cGMP, and IP3.

In a preferred embodiment, a T2R polypeptide is expressed in a eukaryotic cell as a chimeric receptor with a heterologous, chaperone sequence that facilitates its maturation and targeting through the secretory pathway. In a preferred embodiment, the heterologous sequence is a rhodopsin sequence, such as an N-terminal fragment of a rhodopsin. Such chimeric T2R receptors can be expressed in any eukaryotic cell, such as HEK-293 cells. Preferably, the cells comprise a functional G protein, *e.g.*, Gα15, that is capable of coupling the chimeric receptor to an intracellular signaling pathway or to a signaling protein such as phospholipase Cβ. Activation of such chimeric receptors in such cells can be detected using any standard method, such as by detecting changes in intracellular calcium by detecting FURA-2 dependent fluorescence in the cell.

Methods of assaying for modulators of taste transduction include *in vitro* ligand binding assays using T2R polypeptides, portions thereof such as the extracellular domain or transmembrane region or combination thereof, or chimeric proteins comprising one or more domains of a T2R family member; oocyte or tissue culture cell T2R gene expression, or expression of T2R fragments or fusion proteins, such as rhodopsin fusion proteins; transcriptional activation of T2R genes; phosphorylation and dephosphorylation of T2R family members; G-protein binding to GPCRs; ligand binding assays; voltage, membrane potential and conductance changes; ion flux assays; changes in intracellular second messengers such as cGMP, cAMP and inositol triphosphate; changes in intracellular calcium levels; and neurotransmitter release.

Finally, the invention provides methods of detecting T2R nucleic acid and protein expression, allowing investigation of taste transduction regulation and specific identification of taste receptor cells. T2R family members also provide useful nucleic acid probes for paternity and forensic investigations. T2R genes are also useful as a nucleic acid probe for identifying taste receptor cells, such as foliate, fungiform, circumvallate, geschmackstreifen, and epiglottis taste receptor cells, in particular bitter-

taste receptive, gustducin expressing cells. T2R receptors can also be used to generate monoclonal and polyclonal antibodies useful for identifying taste receptor cells. Taste receptor cells can be identified using techniques such as reverse transcription and amplification of mRNA, isolation of total RNA or poly A^+ RNA, northern blotting, dot blotting, *in situ* hybridization, RNase protection, S1 digestion, probing DNA microchip arrays, western blots, and the like.

The T2R genes comprise a large family of related taste cell specific G-protein coupled receptors. Within the genome, these genes are present either alone or within one of several gene clusters. One gene cluster, located at human genomic region 12p13, comprises at least 9 genes, and a second cluster, located at 7q31, comprises at least 4 genes. In total, more than 50 distinct T2R family members have been identified, including several putative pseudogenes. It is estimated that the human genome may contain as many as 80-120 distinct T2R genes, encoding as many as 40-80 functional human receptors.

Some of the T2R genes have been associated with previously mapped mammalian taste-specific loci. For example, the human T2R01 is located at human interval 5p15, precisely where the locus underlying the ability to taste the substance PROP has previously been mapped. In addition, the human gene cluster found at genomic region 12p13 corresponds to a region of mouse chromosome 6 that has been shown to contain numerous bitter-tasting genes, including sucrose octaacetate, ruffinose acetate, cycloheximide, and quinine (*see, e.g.*, Lush *et al.*, Genet. Res. 6:167-174 (1995)). These associations indicate that the T2R genes are involved in the taste detection of various substances, in particular bitter substances. In addition, as shown in Example 7, *infra*, mouse T2R5 is specifically receptive to cycloheximide, and mutations in the mT2R5 gene produce a Cyx phenotype. Similarly, human T2R 4 and mouse T2R8 are specifically receptive to both denatonium and PROP).

Functionally, the T2R genes comprise a family of related seven transmembrane G-protein coupled receptors involved in taste transduction, which interact with a G-protein to mediate taste signal transduction (*see, e.g.*, Fong, *Cell Signal* 8:217 (1996); Baldwin, *Curr. Opin. Cell Biol.* 6:180 (1994)). In particular, T2Rs interact in a ligand-specific manner with the G protein Gustducin.

Structurally, the nucleotide sequence of T2R family members (*see, e.g.*, SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 23, 25, 27, 29, 31, 34, 36, 38, 41, 43, 45, 52, 54, 57, 61, 63, 78, 80, 82, 84,86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110,

112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, [redacted], 140, 142, 144, 146, 148, 150, 152, 154, 156, 157, 159, 161, 163, and 165, isolated from rats, mice, and humans) encodes a family of related polypeptides comprising an extracellular domain, seven transmembrane domains, and a cytoplasmic domain. Related T2R family genes from other species share at least about 60% nucleotide sequence identity over a region of at least about 50 nucleotides in length, optionally 100, 200, 500, or more nucleotides in length, to SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 23, 25, 27, 29, 31, 34, 36, 38, 41, 43, 45, 52, 54, 57, 61, 63, 78, 80, 82, 84,86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 157, 159, 161, 163, or 165, or encode polypeptides sharing at least about 60% amino acid sequence identity over an amino acid region at least about 25 amino acids in length, optionally 50 to 100 amino acids in length to SEQ ID NO:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 22, 24, 26, 28, 30, 32, 33, 35, 37, 39, 40, 42, 44, 46-51, 53, 55, 56, 58-60, 62, 64-77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145, 147, 149, 151, 153, 155, 158, 160, 162, or 164. T2R genes are specifically expressed in taste cells.

Several consensus amino acid sequences or domains have also been identified that are characteristic of T2R family members. For example, T2R family members typically comprise a sequence having at least about 50%, optionally 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or higher, identity to SEQ ID NO:166 (corresponding, *e.g.*, to amino acid positions 16-35 in SEQ ID NO:1, and to T2R transmembrane region 1), SEQ ID NO:167 (corresponding, *e.g.*, to amino acid positions 45-58 in SEQ ID NO:1, and to T2R transmembrane region 2), SEQ ID NO:168 (corresponding, *e.g.*, to amino acid positions 89-101 in SEQ ID NO:1, and to T2R transmembrane region 3), SEQ ID NO:169 (corresponding, *e.g.*, to amino acid positions 102-119 in SEQ ID NO:1, and to T2R transmembrane region 3), SEQ ID NO:170 (corresponding, *e.g.*, to amino acid positions 196-209 in SEQ ID NO:1, and to T2R transmembrane region 5), or SEQ ID NO:171 (corresponding, *e.g.*, to amino acid positions 273-286 in SEQ ID NO:35, and to T2R transmembrane region 7). These conserved domains thus can be used to identify members of the T2R family, by % identity, specific hybridization or amplification, or specific binding by antibodies raised against a domain.

Several T2R genes represent apparent orthologs of each other. For example, human T2R01 (SEQ ID NOs:1, 2), rat T2R01 (SEQ ID NOs:77, 78), and mouse T2R19 (SEQ ID NOs:141, 142), are apparent orthologs. In addition, rat T2R08 (SEQ ID NOs:91, 92) and mouse T2R02 (SEQ ID NOs:107, 108) are about 74% identical at the amino acid sequence level, and are each at least about 50% identical to human T2R13 (SEQ ID NOs:24, 25). Rat T2R03 (SEQ ID NOs:81, 82) and mouse T2R18 (SEQ ID NOs:139, 140) are about 92% identical, and are each at least about 50% identical to human T2R16 (SEQ ID NOs:30, 31). Finally, human T2R04 (SEQ ID NOs:7, 8) and mouse T2R08 (SEQ ID NOs:119, 120) are about 67% identical to each other.

The present invention also provides polymorphic variants of the T2R proteins provided herein. For example, in the rat T2R depicted in SEQ ID NO:77: variant #1, in which an isoleucine residue is substituted for a leucine residue at amino acid position 7; and variant #2, in which an alanine residue is substituted for a glycine residue at amino acid position 20.

The present invention also provides polymorphic variants of the T2R protein depicted in SEQ ID NO:79: variant #1, in which a tyrosine residue is substituted for a phenylalanine residue at amino acid position 2; and variant #2, in which a valine residue is substituted for an isoleucine residue at amino acid position 62.

The present invention also provides polymorphic variants of the T2R protein depicted in SEQ ID NO:81: variant #1, in which a glutamine residue is substituted for an asparagine residue at amino acid position 179; and variant #2, in which a cysteine residue is substituted for a methionine residue at amino acid position 183.

The present invention also provides polymorphic variants of the T2R protein depicted in SEQ ID NO:83: variant #1, in which a glycine residue is substituted for an alanine residue at amino acid position 4; and variant #2, in which a leucine residue is substituted for an isoleucine residue at amino acid position 63.

The present invention also provides polymorphic variants of the T2R protein depicted in SEQ ID NO:85: variant #1, in which a valine residue is substituted for an isoleucine residue at amino acid position 56; and variant #2, in which a methionine residue is substituted for a cysteine residue at amino acid position 57.

The present invention also provides polymorphic variants of the T2R protein depicted in SEQ ID NO:87: variant #1, in which an isoleucine residue is substituted for a valine residue at amino acid position 4; and variant #2, in which an alanine residue is substituted for a glycine residue at amino acid position 5.

The present invention also provides polymorphic variants of the T2R protein depicted in SEQ ID NO:89: variant #1, in which an alanine residue is substituted for a glycine residue at amino acid position 79; and variant #2, in which an arginine residue is substituted for a lysine residue at amino acid position 127.

The present invention also provides polymorphic variants of the T2R protein depicted in SEQ ID NO:91: variant #1, in which a leucine residue is substituted for a valine residue at amino acid position 28; and variant #2, in which an arginine residue is substituted for a lysine residue at amino acid position 80.

The present invention also provides polymorphic variants of the T2R protein depicted in SEQ ID NO:93: variant #1, in which an arginine residue is substituted for a lysine residue at amino acid position 75; and variant #2, in which a methionine residue is substituted for a cysteine residue at amino acid position 251.

The present invention also provides polymorphic variants of the T2R protein depicted in SEQ ID NO:95: variant #1, in which a threonine residue is substituted for a serine residue at amino acid position 48; and variant #2, in which an isoleucine residue is substituted for a valine residue at amino acid position 49.

The present invention also provides polymorphic variants of the T2R protein depicted in SEQ ID NO:97: variant #1, in which a glutamic acid residue is substituted for an aspartic acid residue at amino acid position 25; and variant #2, in which an isoleucine residue is substituted for a leucine residue at amino acid position 100.

The present invention also provides polymorphic variants of the T2R protein depicted in SEQ ID NO:99: variant #1, in which a serine residue is substituted for a threonine residue at amino acid position 4; and variant #2, in which an isoleucine residue is substituted for a valine residue at amino acid position 74.

The present invention also provides polymorphic variants of the T2R protein depicted in SEQ ID NO:101: variant #1, in which an asparagine residue is substituted for a glutamine residue at amino acid position 9; and variant #2, in which a tryptophan residue is substituted for a tyrosine residue at amino acid position 18.

The present invention also provides polymorphic variants of the T2R protein depicted in SEQ ID NO:103: variant #1, in which a threonine residue is substituted for a serine residue at amino acid position 26; and variant #2, in which an isoleucine residue is substituted for a valine residue at amino acid position 8.

The present invention also provides polymorphic variants of the T2R protein depicted in SEQ ID NO:105: variant #1, in which an isoleucine residue is

substituted for a leucine residue at amino acid position 4; and variant #2, in which an arginine residue is substituted for a lysine residue at amino acid position 46.

The present invention also provides polymorphic variants of the T2R protein depicted in SEQ ID NO:107: variant #1, in which a threonine residue is substituted for a serine residue at amino acid position 3; and variant #2, in which an isoleucine residue is substituted for a valine residue at amino acid position 28.

The present invention also provides polymorphic variants of the T2R protein depicted in SEQ ID NO:109: variant #1, in which an isoleucine residue is substituted for a leucine residue at amino acid position 26; and variant #2, in which an arginine residue is substituted for a lysine residue at amino acid position 50.

The present invention also provides polymorphic variants of the T2R protein depicted in SEQ ID NO:111: variant #1, in which a glycine residue is substituted for an alanine residue at amino acid position 4; and variant #2, in which a phenylalanine residue is substituted for a tryptophan residue at amino acid position 60.

The present invention also provides polymorphic variants of the T2R protein depicted in SEQ ID NO:113: variant #1, in which an isoleucine residue is substituted for a leucine residue at amino acid position 62; and variant #2, in which an alanine residue is substituted for a glycine residue at amino acid position 244.

The present invention also provides polymorphic variants of the T2R protein depicted in SEQ ID NO:115: variant #1, in which a serine residue is substituted for a threonine residue at amino acid position 3; and variant #2, in which a lysine residue is substituted for an arginine residue at amino acid position 123.

The present invention also provides polymorphic variants of the T2R protein depicted in SEQ ID NO:117: variant #1, in which an asparagine residue is substituted for a glutamine residue at amino acid position 65; and variant #2, in which a leucine residue is substituted for an isoleucine residue at amino acid position 68.

The present invention also provides polymorphic variants of the T2R protein depicted in SEQ ID NO:119: variant #1, in which an isoleucine residue is substituted for a leucine residue at amino acid position 2; and variant #2, in which an aspartic acid residue is substituted for a glutamic acid residue at amino acid position 4.

The present invention also provides polymorphic variants of the T2R protein depicted in SEQ ID NO:121: variant #1, in which an isoleucine residue is substituted for a leucine residue at amino acid position 16; and variant #2, in which an arginine residue is substituted for a lysine residue at amino acid position 46.

The present invention also provides polymorphic variants of the T2R protein depicted in SEQ ID NO:123: variant #1, in which a threonine residue is substituted for a serine residue at amino acid position 9; and variant #2, in which a tryptophan residue is substituted for a phenylalanine residue at amino acid position 14.

The present invention also provides polymorphic variants of the T2R protein depicted in SEQ ID NO:125: variant #1, in which an isoleucine residue is substituted for a leucine residue at amino acid position 24; and variant #2, in which an arginine residue is substituted for a lysine residue at amino acid position 53.

The present invention also provides polymorphic variants of the T2R protein depicted in SEQ ID NO:127: variant #1, in which a phenylalanine residue is substituted for a tryptophan residue at amino acid position 51; and variant #2, in which an arginine residue is substituted for a lysine residue at amino acid position 101.

The present invention also provides polymorphic variants of the T2R protein depicted in SEQ ID NO:129: variant #1, in which an isoleucine residue is substituted for a valine residue at amino acid position 4; and variant #2, in which a glycine residue is substituted for an alanine residue at amino acid position 52.

The present invention also provides polymorphic variants of the T2R protein depicted in SEQ ID NO:131: variant #1, in which an arginine residue is substituted for a lysine residue at amino acid position 150; and variant #2, in which a leucine residue is substituted for a valine residue at amino acid position 225.

The present invention also provides polymorphic variants of the T2R protein depicted in SEQ ID NO:133: variant #1, in which a leucine residue is substituted for an isoleucine residue at amino acid position 27; and variant #2, in which a lysine residue is substituted for an arginine residue at amino acid position 127.

The present invention also provides polymorphic variants of the T2R protein depicted in SEQ ID NO:135: variant #1, in which a threonine residue is substituted for a serine residue at amino acid position 102; and variant #2, in which a glutamic acid residue is substituted for an aspartic acid residue at amino acid position 220.

The present invention also provides polymorphic variants of the T2R protein depicted in SEQ ID NO:137: variant #1, in which an isoleucine residue is substituted for a leucine residue at amino acid position 24; and variant #2, in which an arginine residue is substituted for a lysine residue at amino acid position 45.

The present invention also provides polymorphic variants of the T2R protein depicted in SEQ ID NO:139: variant #1, in which a leucine residue is substituted for an isoleucine residue at amino acid position 50; and variant #2, in which an alanine residue is substituted for a glycine residue at amino acid position 53.

The present invention also provides polymorphic variants of the T2R protein depicted in SEQ ID NO:141: variant #1, in which a serine residue is substituted for a threonine residue at amino acid position 76; and variant #2, in which an isoleucine residue is substituted for a leucine residue at amino acid position 131.

The present invention also provides polymorphic variants of the T2R protein depicted in SEQ ID NO:143: variant #1, in which an alanine residue is substituted for a glycine residue at amino acid position 98; and variant #2, in which a phenylalanine residue is substituted for a tryptophan residue at amino acid position 153.

The present invention also provides polymorphic variants of the T2R protein depicted in SEQ ID NO:145: variant #1, in which a leucine residue is substituted for an isoleucine residue at amino acid position 8; and variant #2, in which a glycine residue is substituted for an alanine residue at amino acid position 100.

The present invention also provides polymorphic variants of the T2R protein depicted in SEQ ID NO:147: variant #1, in which a glycine residue is substituted for an alanine residue at amino acid position 52; and variant #2, in which a valine residue is substituted for a leucine residue at amino acid position 75.

The present invention also provides polymorphic variants of the T2R protein depicted in SEQ ID NO:149: variant #1, in which a lysine residue is substituted for an arginine residue at amino acid position 44; and variant #2, in which a leucine residue is substituted for a valine residue at amino acid position 49.

The present invention also provides polymorphic variants of the T2R protein depicted in SEQ ID NO:151: variant #1, in which an isoleucine residue is substituted for a leucine residue at amino acid position 5; and variant #2, in which an alanine residue is substituted for a glycine residue at amino acid position 25.

The present invention also provides polymorphic variants of the T2R protein depicted in SEQ ID NO:153: variant #1, in which a glutamic acid residue is substituted for an aspartic acid residue at amino acid position 7; and variant #2, in which an isoleucine residue is substituted for a leucine residue at amino acid position 60.

The present invention also provides polymorphic variants of the T2R protein depicted in SEQ ID NO:155: variant #1, in which an isoleucine residue is

substituted for a valine residue at amino acid position 7; and variant #2, in which a glycine residue is substituted for an alanine residue at amino acid position 23.

The present invention also provides polymorphic variants of the T2R protein depicted in SEQ ID NO:158: variant #1, in which an isoleucine residue is substituted for a leucine residue at amino acid position 5; and variant #2, in which an alanine residue is substituted for a glycine residue at amino acid position 21.

The present invention also provides polymorphic variants of the T2R protein depicted in SEQ ID NO:160: variant #1, in which a leucine residue is substituted for a valine residue at amino acid position 5; and variant #2, in which an alanine residue is substituted for a glycine residue at amino acid position 23.

The present invention also provides polymorphic variants of the T2R protein depicted in SEQ ID NO:162: variant #1, in which an isoleucine residue is substituted for a leucine residue at amino acid position 22; and variant #2, in which an alanine residue is substituted for a glycine residue at amino acid position 34.

The present invention also provides polymorphic variants of the T2R protein depicted in SEQ ID NO:164: variant #1, in which a leucine residue is substituted for an isoleucine residue at amino acid position 49; and variant #2, in which an arginine residue is substituted for a lysine residue at amino acid position 76.

Specific regions of the T2R nucleotide and amino acid sequences may be used to identify polymorphic variants, interspecies homologs, and alleles of T2R family members. This identification can be made *in vitro*, *e.g.*, under stringent hybridization conditions or PCR (*e.g.*, using primers encoding SEQ ID NOS:166-171) and sequencing, or by using the sequence information in a computer system for comparison with other nucleotide sequences. Typically, identification of polymorphic variants and alleles of T2R family members is made by comparing an amino acid sequence of about 25 amino acids or more, *e.g.*, 50-100 amino acids. Amino acid identity of approximately at least 60% or above, optionally 65%, 70%, 75%, 80%, 85%, or 90-95% or above typically demonstrates that a protein is a polymorphic variant, interspecies homolog, or allele of a T2R family member. Sequence comparison can be performed using any of the sequence comparison algorithms discussed below. Antibodies that bind specifically to T2R polypeptides or a conserved region thereof can also be used to identify alleles, interspecies homologs, and polymorphic variants.

Polymorphic variants, interspecies homologs, and alleles of T2R genes are confirmed by examining taste cell specific expression of the putative T2R polypeptide.

Typically, T2R polypeptides having an amino acid sequence of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5; SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:35, SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:44, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:62, SEQ ID NO:64, SEQ ID NO:65, SEQ ID NO:66, SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, SEQ ID NO:76, SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81, SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID NO:158, SEQ ID NO:160, SEQ ID NO:162, or SEQ ID NO:164 is used as a positive control in comparison to the putative T2R protein to demonstrate the identification of a polymorphic variant or allele of the T2R family member. The polymorphic variants, alleles and interspecies homologs are expected to retain the seven transmembrane structure of a G-protein coupled receptor.

The present invention also provides nucleotide sequences for T2R promoters, which can be used to drive taste cell-specific expression of polynucleotides, especially in gustducin expressing taste cells that are receptive to bitter tastants.

Nucleotide and amino acid sequence information for T2R family members may also be used to construct models of taste cell specific polypeptides in a computer system. These models are subsequently used to identify compounds that can activate or inhibit T2R receptor proteins. Such compounds that modulate the activity of T2R family members can be used to investigate the role of T2R genes in taste transduction.

The isolation of T2R family members provides a means for assaying for inhibitors and activators of G-protein coupled receptor taste transduction. Biologically active T2R proteins are useful for testing inhibitors and activators of T2R as taste

transducers, especially bitter taste transducers, using *in vivo* and *in vitro* assays that measure, *e.g.*, transcriptional activation of T2R-dependent genes; ligand binding; phosphorylation and dephosphorylation; binding to G-proteins; G-protein activation; regulatory molecule binding; voltage, membrane potential and conductance changes; ion flux; intracellular second messengers such as cGMP, cAMP and inositol triphosphate; intracellular calcium levels; and neurotransmitter release. Such activators and inhibitors identified using T2R family members can be used to further study taste transduction and to identify specific taste agonists and antagonists. Such activators and inhibitors are useful as pharmaceutical and food agents for customizing taste, for example to decrease the bitter taste of foods or pharmaceuticals.

The present invention also provides assays, preferably high throughput assays, to identify molecules that interact with and/or modulate a T2R polypeptide. In numerous assays, a particular domain of a T2R family member is used, *e.g.*, an extracellular, transmembrane, or intracellular domain or region. In numerous embodiments, an extracellular domain or transmembrane region or combination thereof is bound to a solid substrate, and used, *e.g.*, to isolate ligands, agonists, antagonists, or any other molecule that can bind to and/or modulate the activity of an extracellular domain or transmembrane region of a T2R polypeptide. In certain embodiments, a domain of a T2R polypeptide, *e.g.*, an extracellular, transmembrane, or intracellular domain, is fused to a heterologous polypeptide, thereby forming a chimeric polypeptide, *e.g.*, a chimeric polypeptide with G protein coupled receptor activity. Such chimeric polypeptides are useful, *e.g.*, in assays to identify ligands, agonists, antagonists, or other modulators of a T2R polypeptide. In addition, such chimeric polypeptides are useful to create novel taste receptors with novel ligand binding specificity, modes of regulation, signal transduction pathways, or other such properties, or to create novel taste receptors with novel combinations of ligand binding specificity, modes of regulation, signal transduction pathways, *etc.*

Methods of detecting T2R nucleic acids and expression of T2R polypeptides are also useful for identifying taste cells and creating topological maps of the tongue and the relation of tongue taste receptor cells to taste sensory neurons in the brain. In particular, methods of detecting T2R can be used to identify taste cells sensitive to bitter tastants. Chromosome localization of the genes encoding human T2R genes can be used to identify diseases, mutations, and traits caused by and associated with T2R family members.

II. Definitions

As used herein, the following terms have the meanings ascribed to them unless specified otherwise.

"Taste cells" include neuroepithelial cells that are organized into groups to form taste buds of the tongue, *e.g.*, foliate, fungiform, and circumvallate cells (*see, e.g.*, Roper *et al., Ann. Rev. Neurosci.* 12:329-353 (1989)). Taste cells also include cells of the palate, and other tissues that may contain taste cells such as the esophagus and the stomach.

"T2R" refers to one or more members of a family of G-protein coupled receptors that are expressed in taste cells such as foliate, fungiform, and circumvallate cells, as well as cells of the palate, esophagus, and stomach (*see, e.g.*, Hoon *et al., Cell* 96:541-551 (1999), herein incorporated by reference in its entirety). This family is also referred to as the "SF family" (*see, e.g.*, USSN 09/393,634). Such taste cells can be identified because they express specific molecules such as Gustducin, a taste cell specific G protein, or other taste specific molecules (McLaughin *et al., Nature* 357:563-569 (1992)). Taste receptor cells can also be identified on the basis of morphology (*see, e.g.*, Roper, *supra*). T2R family members have the ability to act as receptors for taste transduction. T2R family members are also referred to as the "GR" family, for gustatory receptor, or "SF" family.

"T2R" nucleic acids encode a family of GPCRs with seven transmembrane regions that have "G-protein coupled receptor activity," *e.g.*, they bind to G-proteins in response to extracellular stimuli and promote production of second messengers such as IP3, cAMP, cGMP, and Ca^{2+} via stimulation of enzymes such as phospholipase C and adenylate cyclase (for a description of the structure and function of GPCRs, *see, e.g.*, Fong, *supra*, and Baldwin, *supra*). A dendogram providing the relationship between certain T2R family members is provided as Figure 2. These nucleic acids encode proteins that are expressed in taste cells, in particular Gustducin-expressing taste cells that are responsive to bitter tastants. A single taste cell may contain many distinct T2R polypeptides.

The term "T2R" family therefore refers to polymorphic variants, alleles, mutants, and interspecies homologs that: (1) have about 60% amino acid sequence identity, optionally about 75, 80, 85, 90, or 95% amino acid sequence identity to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5; SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ

ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:35, SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:44, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:62, SEQ ID NO:64, SEQ ID NO:65, SEQ ID NO:66, SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, SEQ ID NO:76, SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81, SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID NO:158, SEQ ID NO:160, SEQ ID NO:162, or SEQ ID NO:164 over a window of about 25 amino acids, optionally 50-100 amino acids; (2) specifically bind to antibodies raised against an immunogen comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5; SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:35, SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:44, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:62, SEQ ID NO:64, SEQ ID NO:65, SEQ ID NO:66, SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, SEQ ID NO:76, SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81, SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID

NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID NO:158, SEQ ID NO:160, SEQ ID NO:162, and SEQ ID NO:164, and conservatively modified variants thereof; (3) specifically hybridize (with a size of at least about 100, optionally at least about 500-1000 nucleotides) under stringent hybridization conditions to a sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:20, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:34, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:41, SEQ ID NO:43, SEQ ID NO:45, SEQ ID NO:52, SEQ ID NO:54, SEQ ID NO:57, SEQ ID NO:61, SEQ ID NO:63, SEQ ID NO:78, SEQ ID NO:80, SEQ ID NO:82, SEQ ID NO:84, SEQ ID NO:86; SEQ ID NO:88, SEQ ID NO:90, SEQ ID NO:92, SEQ ID NO:94, SEQ ID NO:96, SEQ ID NO:98, SEQ ID NO:100, SEQ ID NO:102, SEQ ID NO:104 SEQ ID NO:106, SEQ ID NO:108, SEQ ID NO:110, SEQ ID NO:112, SEQ ID NO:114, SEQ ID NO:116, SEQ ID NO:118, SEQ ID NO:120, SEQ ID NO:120, SEQ ID NO:122, SEQ ID NO:124, SEQ ID NO:126, SEQ ID NO:128, SEQ ID NO:130, SEQ ID NO:132, SEQ ID NO:134, SEQ ID NO:136, SEQ ID NO:138, SEQ ID NO:140, SEQ ID NO:142, SEQ ID NO:144, SEQ ID NO:146, SEQ ID NO:148, SEQ ID NO:150, SEQ ID NO:152, SEQ ID NO:154, SEQ ID NO:156, SEQ ID NO:157, SEQ ID NO:159, SEQ ID NO:161, SEQ ID NO:163, and SEQ ID NO:165, and conservatively modified variants thereof; (4) comprise a sequence at least about 50% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:166, SEQ ID NO:167, SEQ ID NO:168, SEQ ID NO:169, SEQ ID NO:170, and SEQ ID NO:171; or (5) are amplified by primers that specifically hybridize under stringent hybridization conditions to the same sequence as a degenerate primer sets encoding SEQ ID NO:166, SEQ ID NO:167, SEQ ID NO:168, SEQ ID NO:169, SEQ ID NO:170, or SEQ ID NO:171.

Topologically, sensory GPCRs have an "N-terminal domain" "extracellular domains," a "transmembrane domain" comprising seven transmembrane regions, cytoplasmic, and extracellular loops, "cytoplasmic domains," and a "C-terminal domain" (*see, e.g.*, Hoon *et al.*, *Cell* 96:541-551 (1999); Buck & Axel, *Cell* 65:175-187 (1991)). These domains can be structurally identified using methods known to those of

skill in the art, such as sequence analysis programs that identify hydrophobic and hydrophilic domains (*see, e.g.*, Stryer, *Biochemistry* (3rd ed. 1988); *see also* any of a number of Internet based sequence analysis programs, such as those found at dot.imgen.bcm.tmc.edu). Such domains are useful for making chimeric proteins and for *in vitro* assays of the invention, *e.g.*, ligand binding assays.

"Extracellular domains" therefore refers to the domains of T2R polypeptides that protrude from the cellular membrane and are exposed to the extracellular face of the cell. Such domains would include the "N terminal domain" that is exposed to the extracellular face of the cell, as well as the extracellular loops of the transmembrane domain that are exposed to the extracellular face of the cell, *i.e.*, the loops between transmembrane regions 2 and 3, and between transmembrane regions 4 and 5. The "N terminal domain" region starts at the N-terminus and extends to a region close to the start of the transmembrane domain. These extracellular domains are useful for *in vitro* ligand binding assays, both soluble and solid phase. In addition, transmembrane regions, described below, can also bind ligand either in combination with the extracellular domain or alone, and are therefore also useful for *in vitro* ligand binding assays.

"Transmembrane domain," which comprises the seven transmembrane "regions," refers to the domain of T2R polypeptides that lies within the plasma membrane, and may also include the corresponding cytoplasmic (intracellular) and extracellular loops, also referred to as transmembrane domain "regions." The seven transmembrane regions and extracellular and cytoplasmic loops can be identified using standard methods, as described in Kyte & Doolittle, *J. Mol. Biol.* 157:105-132 (1982)), or in Stryer, *supra*.

"Cytoplasmic domains" refers to the domains of T2R proteins that face the inside of the cell, e.g., the "C terminal domain" and the intracellular loops of the transmembrane domain, e.g., the intracellular loops between transmembrane regions 1 and 2, and the intracellular loops between transmembrane regions 3 and 4. "C terminal domain" refers to the region that spans the end of the last transmembrane domain and the C-terminus of the protein, and which is normally located within the cytoplasm.

"Biological sample" as used herein is a sample of biological tissue or fluid that contains one or more T2R nucleic acids encoding one or more T2R proteins. Such samples include, but are not limited to, tissue isolated from humans, mice, and rats, in particular, tongue, palate, and other tissues that may contain taste cells such as the esophagus and the stomach. Biological samples may also include sections of tissues such

as frozen sections taken for histological purposes. A biological sample is typically obtained from a eukaryotic organism, such as insects, protozoa, birds, fish, reptiles, and preferably a mammal such as rat, mouse, cow, dog, guinea pig, or rabbit, and most preferably a primate such as chimpanzees or humans.

"GPCR activity" refers to the ability of a GPCR to transduce a signal. Such activity can be measured in a heterologous cell, by coupling a GPCR (or a chimeric GPCR) to either a G-protein or promiscuous G-protein such as Gα15, and an enzyme such as PLC, and measuring increases in intracellular calcium using (Offermans & Simon, *J. Biol. Chem.* 270:15175-15180 (1995)). Receptor activity can be effectively measured by recording ligand-induced changes in $[Ca^{2+}]_i$ using fluorescent Ca^{2+}-indicator dyes and fluorometric imaging. Optionally, the polypeptides of the invention are involved in sensory transduction, optionally taste transduction in taste cells.

The phrase "functional effects" in the context of assays for testing compounds that modulate T2R family member mediated taste transduction includes the determination of any parameter that is indirectly or directly under the influence of the receptor, *e.g.*, functional, physical and chemical effects. It includes ligand binding, changes in ion flux, membrane potential, current flow, transcription, G-protein binding, GPCR phosphorylation or dephosphorylation, signal transduction, receptor-ligand interactions, second messenger concentrations (*e.g.*, cAMP, cGMP, IP3, or intracellular Ca^{2+}), *in vitro*, *in vivo*, and *ex vivo* and also includes other physiologic effects such increases or decreases of neurotransmitter or hormone release.

By "determining the functional effect" is meant assays for a compound that increases or decreases a parameter that is indirectly or directly under the influence of a T2R family member, *e.g.*, functional, physical and chemical effects. Such functional effects can be measured by any means known to those skilled in the art, *e.g.*, changes in spectroscopic characteristics (*e.g.*, fluorescence, absorbance, refractive index), hydrodynamic (*e.g.*, shape), chromatographic, or solubility properties, patch clamping, voltage-sensitive dyes, whole cell currents, radioisotope efflux, inducible markers, oocyte T2R gene expression; tissue culture cell T2R expression; transcriptional activation of T2R genes; ligand binding assays; voltage, membrane potential and conductance changes; ion flux assays; changes in intracellular second messengers such as cAMP, cGMP, and inositol triphosphate (IP3); changes in intracellular calcium levels; neurotransmitter release, and the like.

"Inhibitors," "activators," and "modulators" of T2R genes or proteins are used interchangeably to refer to inhibitory, activating, or modulating molecules identified using *in vitro* and *in vivo* assays for taste transduction, *e.g.*, ligands, agonists, antagonists, and their homologs and mimetics. Inhibitors are compounds that, *e.g.*, bind to, partially or totally block stimulation, decrease, prevent, delay activation, inactivate, desensitize, or down regulate taste transduction, *e.g.*, antagonists. Activators are compounds that, *e.g.*, bind to, stimulate, increase, open, activate, facilitate, enhance activation, sensitize or up regulate taste transduction, *e.g.*, agonists. Modulators include compounds that, *e.g.*, alter the interaction of a receptor with: extracellular proteins that bind activators or inhibitor (*e.g.*, ebnerin and other members of the hydrophobic carrier family); G -proteins; kinases (*e.g.*, homologs of rhodopsin kinase and beta adrenergic receptor kinases that are involved in deactivation and desensitization of a receptor); and arrestin-like proteins, which also deactivate and desensitize receptors. Modulators include genetically modified versions of T2R family members, *e.g.*, with altered activity, as well as naturally occurring and synthetic ligands, antagonists, agonists, small chemical molecules and the like. Such assays for inhibitors and activators include, *e.g.*, expressing T2R family members in cells or cell membranes, applying putative modulator compounds, in the presence or absence of tastants, *e.g.*, bitter tastants, and then determining the functional effects on taste transduction, as described above. Samples or assays comprising T2R family members that are treated with a potential activator, inhibitor, or modulator are compared to control samples without the inhibitor, activator, or modulator to examine the extent of inhibition. Control samples (untreated with inhibitors) are assigned a relative T2R activity value of 100%. Inhibition of a T2R is achieved when the T2R activity value relative to the control is about 80%, optionally 50% or 25-0%. Activation of a T2R is achieved when the T2R activity value relative to the control is 110%, optionally 150%, optionally 200-500%, or 1000-3000% higher.

"Biologically active" T2R refers to a T2R having GPCR activity as described above, involved in taste transduction in taste receptor cells, in particular bitter taste transduction.

The terms "isolated" "purified" or "biologically pure" refer to material that is substantially or essentially free from components which normally accompany it as found in its native state. Purity and homogeneity are typically determined using analytical chemistry techniques such as polyacrylamide gel electrophoresis or high performance liquid chromatography. A protein that is the predominant species present in

a preparation is substantially purified. In particular, an isolated T2R nucleic acid is separated from open reading frames that flank the T2R gene and encode proteins other than a T2R. The term "purified" denotes that a nucleic acid or protein gives rise to essentially one band in an electrophoretic gel. Particularly, it means that the nucleic acid or protein is at least 85% pure, optionally at least 95% pure, and optionally at least 99% pure.

"Nucleic acid" refers to deoxyribonucleotides or ribonucleotides and polymers thereof in either single- or double-stranded form. The term encompasses nucleic acids containing known nucleotide analogs or modified backbone residues or linkages, which are synthetic, naturally occurring, and non-naturally occurring, which have similar binding properties as the reference nucleic acid, and which are metabolized in a manner similar to the reference nucleotides. Examples of such analogs include, without limitation, phosphorothioates, phosphoramidates, methyl phosphonates, chiral-methyl phosphonates, 2-O-methyl ribonucleotides, peptide-nucleic acids (PNAs).

Unless otherwise indicated, a particular nucleic acid sequence also implicitly encompasses conservatively modified variants thereof (*e.g.*, degenerate codon substitutions) and complementary sequences, as well as the sequence explicitly indicated. Specifically, degenerate codon substitutions may be achieved by generating sequences in which the third position of one or more selected (or all) codons is substituted with mixed-base and/or deoxyinosine residues (Batzer *et al.*, *Nucleic Acid Res.* 19:5081 (1991); Ohtsuka *et al.*, *J. Biol. Chem.* 260:2605-2608 (1985); Rossolini *et al.*, *Mol. Cell. Probes* 8:91-98 (1994)). The term nucleic acid is used interchangeably with gene, cDNA, mRNA, oligonucleotide, and polynucleotide.

The terms "polypeptide," "peptide" and "protein" are used interchangeably herein to refer to a polymer of amino acid residues. The terms apply to amino acid polymers in which one or more amino acid residue is an artificial chemical mimetic of a corresponding naturally occurring amino acid, as well as to naturally occurring amino acid polymers and non-naturally occurring amino acid polymer.

The term "amino acid" refers to naturally occurring and synthetic amino acids, as well as amino acid analogs and amino acid mimetics that function in a manner similar to the naturally occurring amino acids. Naturally occurring amino acids are those encoded by the genetic code, as well as those amino acids that are later modified, *e.g.*, hydroxyproline, γ-carboxyglutamate, and O-phosphoserine. Amino acid analogs refers to compounds that have the same basic chemical structure as a naturally occurring amino

acid, *i.e.*, an α carbon that is bound to a hydrogen, a carboxyl group, an amino group, and an R group, *e.g.*, homoserine, norleucine, methionine sulfoxide, methionine methyl sulfonium. Such analogs have modified R groups (*e.g.*, norleucine) or modified peptide backbones, but retain the same basic chemical structure as a naturally occurring amino acid. Amino acid mimetics refers to chemical compounds that have a structure that is different from the general chemical structure of an amino acid, but that functions in a manner similar to a naturally occurring amino acid.

Amino acids may be referred to herein by either their commonly known three letter symbols or by the one-letter symbols recommended by the IUPAC-IUB Biochemical Nomenclature Commission. Nucleotides, likewise, may be referred to by their commonly accepted single-letter codes.

"Conservatively modified variants" applies to both amino acid and nucleic acid sequences. With respect to particular nucleic acid sequences, conservatively modified variants refers to those nucleic acids which encode identical or essentially identical amino acid sequences, or where the nucleic acid does not encode an amino acid sequence, to essentially identical sequences. Because of the degeneracy of the genetic code, a large number of functionally identical nucleic acids encode any given protein. For instance, the codons GCA, GCC, GCG and GCU all encode the amino acid alanine. Thus, at every position where an alanine is specified by a codon, the codon can be altered to any of the corresponding codons described without altering the encoded polypeptide. Such nucleic acid variations are "silent variations," which are one species of conservatively modified variations. Every nucleic acid sequence herein which encodes a polypeptide also describes every possible silent variation of the nucleic acid. One of skill will recognize that each codon in a nucleic acid (except AUG, which is ordinarily the only codon for methionine, and TGG, which is ordinarily the only codon for tryptophan) can be modified to yield a functionally identical molecule. Accordingly, each silent variation of a nucleic acid which encodes a polypeptide is implicit in each described sequence.

As to amino acid sequences, one of skill will recognize that individual substitutions, deletions or additions to a nucleic acid, peptide, polypeptide, or protein sequence which alters, adds or deletes a single amino acid or a small percentage of amino acids in the encoded sequence is a "conservatively modified variant" where the alteration results in the substitution of an amino acid with a chemically similar amino acid. Conservative substitution tables providing functionally similar amino acids are well

known in the art. S[redacted] conservatively modified variants are in add[redacted]n to and do not exclude polymorphic variants, interspecies homologs, and alleles of the invention.

The following eight groups each contain amino acids that are conservative substitutions for one another:

1) Alanine (A), Glycine (G);
2) Aspartic acid (D), Glutamic acid (E);
3) Asparagine (N), Glutamine (Q);
4) Arginine (R), Lysine (K);
5) Isoleucine (I), Leucine (L), Methionine (M), Valine (V);
6) Phenylalanine (F), Tyrosine (Y), Tryptophan (W);
7) Serine (S), Threonine (T); and
8) Cysteine (C), Methionine (M)

(*see, e.g.*, Creighton, *Proteins* (1984)).

Macromolecular structures such as polypeptide structures can be described in terms of various levels of organization. For a general discussion of this organization, *see, e.g.*, Alberts *et al.*, *Molecular Biology of the Cell* (3rd ed., 1994) and Cantor and Schimmel, *Biophysical Chemistry Part I: The Conformation of Biological Macromolecules* (1980). "Primary structure" refers to the amino acid sequence of a particular peptide. "Secondary structure" refers to locally ordered, three dimensional structures within a polypeptide. These structures are commonly known as domains. Domains are portions of a polypeptide that form a compact unit of the polypeptide and are typically 50 to 350 amino acids long. Typical domains are made up of sections of lesser organization such as stretches of β-sheet and α-helices. "Tertiary structure" refers to the complete three dimensional structure of a polypeptide monomer. "Quaternary structure" refers to the three dimensional structure formed by the noncovalent association of independent tertiary units. Anisotropic terms are also known as energy terms.

A "label" or a "detectable moiety" is a composition detectable by spectroscopic, photochemical, biochemical, immunochemical, or chemical means. For example, useful labels include ^{32}P, fluorescent dyes, electron-dense reagents, enzymes (*e.g.*, as commonly used in an ELISA), biotin, digoxigenin, or haptens and proteins which can be made detectable, *e.g.*, by incorporating a radiolabel into the peptide or used to detect antibodies specifically reactive with the peptide.

A "labeled nucleic acid probe or oligonucleotide" is one that is bound, either covalently, through a linker or a chemical bond, or noncovalently, through ionic, van der Waals, electrostatic, or hydrogen bonds to a label such that the presence of the probe may be detected by detecting the presence of the label bound to the probe.

As used herein a "nucleic acid probe or oligonucleotide" is defined as a nucleic acid capable of binding to a target nucleic acid of complementary sequence through one or more types of chemical bonds, usually through complementary base pairing, usually through hydrogen bond formation. As used herein, a probe may include natural (*i.e.*, A, G, C, or T) or modified bases (7-deazaguanosine, inosine, *etc.*). In addition, the bases in a probe may be joined by a linkage other than a phosphodiester bond, so long as it does not interfere with hybridization. Thus, for example, probes may be peptide nucleic acids in which the constituent bases are joined by peptide bonds rather than phosphodiester linkages. It will be understood by one of skill in the art that probes may bind target sequences lacking complete complementarity with the probe sequence depending upon the stringency of the hybridization conditions. The probes are optionally directly labeled as with isotopes, chromophores, lumiphores, chromogens, or indirectly labeled such as with biotin to which a streptavidin complex may later bind. By assaying for the presence or absence of the probe, one can detect the presence or absence of the select sequence or subsequence.

The term "recombinant" when used with reference, *e.g.*, to a cell, or nucleic acid, protein, or vector, indicates that the cell, nucleic acid, protein or vector, has been modified by the introduction of a heterologous nucleic acid or protein or the alteration of a native nucleic acid or protein, or that the cell is derived from a cell so modified. Thus, for example, recombinant cells express genes that are not found within the native (non-recombinant) form of the cell or express native genes that are otherwise abnormally expressed, under expressed or not expressed at all.

The term "heterologous" when used with reference to portions of a nucleic acid indicates that the nucleic acid comprises two or more subsequences that are not found in the same relationship to each other in nature. For instance, the nucleic acid is typically recombinantly produced, having two or more sequences from unrelated genes arranged to make a new functional nucleic acid, *e.g.*, a promoter from one source and a coding region from another source. Similarly, a heterologous protein indicates that the protein comprises two or more subsequences that are not found in the same relationship to each other in nature (e.g., a fusion protein).

A "promoter" is defined as an array of nucleic acid control sequences that direct transcription of a nucleic acid. As used herein, a promoter includes necessary nucleic acid sequences near the start site of transcription, such as, in the case of a polymerase II type promoter, a TATA element. A promoter also optionally includes distal enhancer or repressor elements, which can be located as much as several thousand base pairs from the start site of transcription. A "constitutive" promoter is a promoter that is active under most environmental and developmental conditions. An "inducible" promoter is a promoter that is active under environmental or developmental regulation. The term "operably linked" refers to a functional linkage between a nucleic acid expression control sequence (such as a promoter, or array of transcription factor binding sites) and a second nucleic acid sequence, wherein the expression control sequence directs transcription of the nucleic acid corresponding to the second sequence.

An "expression vector" is a nucleic acid construct, generated recombinantly or synthetically, with a series of specified nucleic acid elements that permit transcription of a particular nucleic acid in a host cell. The expression vector can be part of a plasmid, virus, or nucleic acid fragment. Typically, the expression vector includes a nucleic acid to be transcribed operably linked to a promoter.

The terms "identical" or percent "identity," in the context of two or more nucleic acids or polypeptide sequences, refer to two or more sequences or subsequences or domains that are the same or have a specified percentage of amino acid residues or nucleotides that are the same (*i.e.*, 50% identity, optionally 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% or higher identity over a specified region), when compared and aligned for maximum correspondence over a comparison window, or designated region as measured using one of the following sequence comparison algorithms or by manual alignment and visual inspection. Such sequences are then said to be "substantially identical." This definition also refers to the compliment of a test sequence. Optionally, the identity exists over a region that is at least about 50 amino acids or nucleotides in length, or more preferably over a region that is 75-100 amino acids or nucleotides in length.

For sequence comparison, typically one sequence acts as a reference sequence, to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are entered into a computer, subsequence coordinates are designated, if necessary, and sequence algorithm program parameters are designated. Default program parameters can be used, as described below for the

BLASTN and BLASTP programs, or alternative parameters can be designated. The sequence comparison algorithm then calculates the percent sequence identities for the test sequences relative to the reference sequence, based on the program parameters.

A "comparison window", as used herein, includes reference to a segment of any one of the number of contiguous positions selected from the group consisting of from 20 to 600, usually about 50 to about 200, more usually about 100 to about 150 in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned. Methods of alignment of sequences for comparison are well-known in the art. Optimal alignment of sequences for comparison can be conducted, e.g., by the local homology algorithm of Smith & Waterman, *Adv. Appl. Math.* 2:482 (1981), by the homology alignment algorithm of Needleman & Wunsch, *J. Mol. Biol.* 48:443 (1970), by the search for similarity method of Pearson & Lipman, *Proc. Nat'l. Acad. Sci. USA* 85:2444 (1988), by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), or by manual alignment and visual inspection (*see, e.g., Current Protocols in Molecular Biology* (Ausubel *et al.*, eds. 1995 supplement)).

A preferred example of an algorithm that is suitable for determining percent sequence identity and sequence similarity are the BLAST and BLAST 2.0 algorithms, which are described in Altschul *et al., Nuc. Acids Res.* 25:3389-3402 (1977) and Altschul *et al., J. Mol. Biol.* 215:403-410 (1990), respectively. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/). This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul *et al., supra*). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always > 0) and N (penalty score for mismatching residues; always < 0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by

the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) or 10, M=5, N=-4 and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength of 3, and expectation (E) of 10, and the BLOSUM62 scoring matrix (*see* Henikoff & Henikoff, *Proc. Natl. Acad. Sci. USA* 89:10915 (1989)) alignments (B) of 50, expectation (E) of 10, M=5, N=-4, and a comparison of both strands.

Another example of a useful algorithm is PILEUP. PILEUP creates a multiple sequence alignment from a group of related sequences using progressive, pairwise alignments to show relationship and percent sequence identity. It also plots a tree or dendogram showing the clustering relationships used to create the alignment (*see, e.g.*, Figure 2). PILEUP uses a simplification of the progressive alignment method of Feng & Doolittle, *J. Mol. Evol.* 35:351-360 (1987). The method used is similar to the method described by Higgins & Sharp, *CABIOS* 5:151-153 (1989). The program can align up to 300 sequences, each of a maximum length of 5,000 nucleotides or amino acids. The multiple alignment procedure begins with the pairwise alignment of the two most similar sequences, producing a cluster of two aligned sequences. This cluster is then aligned to the next most related sequence or cluster of aligned sequences. Two clusters of sequences are aligned by a simple extension of the pairwise alignment of two individual sequences. The final alignment is achieved by a series of progressive, pairwise alignments. The program is run by designating specific sequences and their amino acid or nucleotide coordinates for regions of sequence comparison and by designating the program parameters. Using PILEUP, a reference sequence is compared to other test sequences to determine the percent sequence identity relationship using the following parameters: default gap weight (3.00), default gap length weight (0.10), and weighted end gaps. PILEUP can be obtained from the GCG sequence analysis software package, *e.g.*, version 7.0 (Devereaux *et al., Nuc. Acids Res.* 12:387-395 (1984)).

An indication that two nucleic acid sequences or polypeptides are substantially identical is that the polypeptide encoded by the first nucleic acid is immunologically cross reactive with the antibodies raised against the polypeptide encoded by the second nucleic acid, as described below. Thus, a polypeptide is typically

substantially identi[redacted]to a second polypeptide, for example, wher[redacted]e two peptides differ only by conservative substitutions. Another indication that two nucleic acid sequences are substantially identical is that the two molecules or their complements hybridize to each other under stringent conditions, as described below. Yet another indication that two nucleic acid sequences are substantially identical is that the same primers can be used to amplify the sequence.

The phrase "selectively (or specifically) hybridizes to" refers to the binding, duplexing, or hybridizing of a molecule only to a particular nucleotide sequence under stringent hybridization conditions when that sequence is present in a complex mixture (*e.g.*, total cellular or library DNA or RNA).

The phrase "stringent hybridization conditions" refers to conditions under which a probe will hybridize to its target subsequence, typically in a complex mixture of nucleic acid, but to no other sequences. Stringent conditions are sequence-dependent and will be different in different circumstances. Longer sequences hybridize specifically at higher temperatures. An extensive guide to the hybridization of nucleic acids is found in Tijssen, *Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Probes*, "Overview of principles of hybridization and the strategy of nucleic acid assays" (1993). Generally, stringent conditions are selected to be about 5-10° C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength pH. The T_m is the temperature (under defined ionic strength, pH, and nucleic concentration) at which 50% of the probes complementary to the target hybridize to the target sequence at equilibrium (as the target sequences are present in excess, at T_m, 50% of the probes are occupied at equilibrium). Stringent conditions will be those in which the salt concentration is less than about 1.0 M sodium ion, typically about 0.01 to 1.0 M sodium ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30° C for short probes (*e.g.*, 10 to 50 nucleotides) and at least about 60° C for long probes (*e.g.*, greater than 50 nucleotides). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide. For selective or specific hybridization, a positive signal is at least two times background, optionally 10 times background hybridization. Exemplary stringent hybridization conditions can be as following: 50% formamide, 5x SSC, and 1% SDS, incubating at 42°C, or, 5x SSC, 1% SDS, incubating at 65°C, with wash in 0.2x SSC, and 0.1% SDS at 65°C. Such hybridizations and wash steps can be carried out for, *e.g.*, 1, 2, 5, 10, 15, 30, 60, or more minutes.

Nucleic acids that do not hybridize to each other under stringent conditions are still substantially identical if the polypeptides which they encode are substantially identical. This occurs, for example, when a copy of a nucleic acid is created using the maximum codon degeneracy permitted by the genetic code. In such cases, the nucleic acids typically hybridize under moderately stringent hybridization conditions. Exemplary "moderately stringent hybridization conditions" include a hybridization in a buffer of 40% formamide, 1 M NaCl, 1% SDS at 37°C, and a wash in 1X SSC at 45°C. Such hybridizations and wash steps can be carried out for, *e.g.*, 1, 2, 5, 10, 15, 30, 60, or more minutes. A positive hybridization is at least twice background. Those of ordinary skill will readily recognize that alternative hybridization and wash conditions can be utilized to provide conditions of similar stringency.

"Antibody" refers to a polypeptide comprising a framework region from an immunoglobulin gene or fragments thereof that specifically binds and recognizes an antigen. The recognized immunoglobulin genes include the kappa, lambda, alpha, gamma, delta, epsilon, and mu constant region genes, as well as the myriad immunoglobulin variable region genes. Light chains are classified as either kappa or lambda. Heavy chains are classified as gamma, mu, alpha, delta, or epsilon, which in turn define the immunoglobulin classes, IgG, IgM, IgA, IgD and IgE, respectively.

An exemplary immunoglobulin (antibody) structural unit comprises a tetramer. Each tetramer is composed of two identical pairs of polypeptide chains, each pair having one "light" (about 25 kDa) and one "heavy" chain (about 50-70 kDa). The N-terminus of each chain defines a variable region of about 100 to 110 or more amino acids primarily responsible for antigen recognition. The terms variable light chain (V_L) and variable heavy chain (V_H) refer to these light and heavy chains respectively.

Antibodies exist, e.g., as intact immunoglobulins or as a number of well-characterized fragments produced by digestion with various peptidases. Thus, for example, pepsin digests an antibody below the disulfide linkages in the hinge region to produce F(ab)'$_2$, a dimer of Fab which itself is a light chain joined to V_H-C_H1 by a disulfide bond. The F(ab)'$_2$ may be reduced under mild conditions to break the disulfide linkage in the hinge region, thereby converting the F(ab)'$_2$ dimer into an Fab' monomer. The Fab' monomer is essentially Fab with part of the hinge region (*see Fundamental Immunology* (Paul ed., 3d ed. 1993). While various antibody fragments are defined in terms of the digestion of an intact antibody, one of skill will appreciate that such fragments may be synthesized *de novo* either chemically or by using recombinant DNA

methodology. Thus, the term antibody, as used herein, also includes antibody fragments either produced by the modification of whole antibodies, or those synthesized *de novo* using recombinant DNA methodologies (*e.g.*, single chain Fv) or those identified using phage display libraries (*see, e.g.*, McCafferty *et al.*, *Nature* 348:552-554 (1990)).

For preparation of monoclonal or polyclonal antibodies, any technique known in the art can be used (*see, e.g.*, Kohler & Milstein, *Nature* 256:495-497 (1975); Kozbor *et al.*, *Immunology Today* 4: 72 (1983); Cole *et al.*, pp. 77-96 in *Monoclonal Antibodies and Cancer Therapy* (1985)). Techniques for the production of single chain antibodies (U.S. Patent 4,946,778) can be adapted to produce antibodies to polypeptides of this invention. Also, transgenic mice, or other organisms such as other mammals, may be used to express humanized antibodies. Alternatively, phage display technology can be used to identify antibodies and heteromeric Fab fragments that specifically bind to selected antigens (*see, e.g.*, McCafferty *et al.*, *Nature* 348:552-554 (1990); Marks *et al.*, *Biotechnology* 10:779-783 (1992)).

A "chimeric antibody" is an antibody molecule in which (a) the constant region, or a portion thereof, is altered, replaced or exchanged so that the antigen binding site (variable region) is linked to a constant region of a different or altered class, effector function and/or species, or an entirely different molecule which confers new properties to the chimeric antibody, *e.g.*, an enzyme, toxin, hormone, growth factor, drug, etc.; or (b) the variable region, or a portion thereof, is altered, replaced or exchanged with a variable region having a different or altered antigen specificity.

An "anti-T2R" antibody is an antibody or antibody fragment that specifically binds a polypeptide encoded by a T2R gene, cDNA, or a subsequence thereof.

The term "immunoassay" is an assay that uses an antibody to specifically bind an antigen. The immunoassay is characterized by the use of specific binding properties of a particular antibody to isolate, target, and/or quantify the antigen.

The phrase "specifically (or selectively) binds" to an antibody or "specifically (or selectively) immunoreactive with," when referring to a protein or peptide, refers to a binding reaction that is determinative of the presence of the protein in a heterogeneous population of proteins and other biologics. Thus, under designated immunoassay conditions, the specified antibodies bind to a particular protein at least two times the background and do not substantially bind in a significant amount to other proteins present in the sample. Specific binding to an antibody under such conditions

may require an antibody that is selected for its specificity for a particular protein. For example, polyclonal antibodies raised to a T2R family member from specific species such as rat, mouse, or human can be selected to obtain only those polyclonal antibodies that are specifically immunoreactive with the T2R protein or an immunogenic portion thereof and not with other proteins, except for orthologs or polymorphic variants and alleles of the T2R protein. This selection may be achieved by subtracting out antibodies that cross-react with T2R molecules from other species or other T2R molecules. Antibodies can also be selected that recognize only T2R GPCR family members but not GPCRs from other families. A variety of immunoassay formats may be used to select antibodies specifically immunoreactive with a particular protein. For example, solid-phase ELISA immunoassays are routinely used to select antibodies specifically immunoreactive with a protein (*see, e.g.*, Harlow & Lane, *Antibodies, A Laboratory Manual* (1988), for a description of immunoassay formats and conditions that can be used to determine specific immunoreactivity). Typically a specific or selective reaction will be at least twice background signal or noise and more typically more than 10 to 100 times background.

In one embodiment, immunogenic domains corresponding to SEQ ID NOs:166-171 can be used to raise antibodies that specifically bind to polypeptides of the T2R family.

The phrase "selectively associates with" refers to the ability of a nucleic acid to "selectively hybridize" with another as defined above, or the ability of an antibody to "selectively (or specifically) bind to a protein, as defined above.

By "host cell" is meant a cell that contains an expression vector and supports the replication or expression of the expression vector. Host cells may be prokaryotic cells such as *E. coli*, or eukaryotic cells such as yeast, insect, amphibian, or mammalian cells such as CHO, HeLa, HEK-293, and the like, *e.g.*, cultured cells, explants, and cells *in vivo*.

III. Isolation of nucleic acids encoding T2R family members

A. General recombinant DNA methods

This invention relies on routine techniques in the field of recombinant genetics. Basic texts disclosing the general methods of use in this invention include Sambrook *et al.*, *Molecular Cloning, A Laboratory Manual* (2nd ed. 1989); Kriegler,

Gene Transfer and Expression: A Laboratory Manual (1990); and *Current Protocols in Molecular Biology* (Ausubel *et al.*, eds., 1994)).

For nucleic acids, sizes are given in either kilobases (kb) or base pairs (bp). These are estimates derived from agarose or acrylamide gel electrophoresis, from sequenced nucleic acids, or from published DNA sequences. For proteins, sizes are given in kilodaltons (kDa) or amino acid residue numbers. Proteins sizes are estimated from gel electrophoresis, from sequenced proteins, from derived amino acid sequences, or from published protein sequences.

Oligonucleotides that are not commercially available can be chemically synthesized according to the solid phase phosphoramidite triester method first described by Beaucage & Caruthers, *Tetrahedron Letts.* 22:1859-1862 (1981), using an automated synthesizer, as described in Van Devanter *et al.*, *Nucleic Acids Res.* 12:6159-6168 (1984). Purification of oligonucleotides is by either native acrylamide gel electrophoresis or by anion-exchange HPLC as described in Pearson & Reanier, *J. Chrom.* 255:137-149 (1983).

The sequence of the cloned genes and synthetic oligonucleotides can be verified after cloning using, *e.g.*, the chain termination method for sequencing double-stranded templates of Wallace *et al.*, *Gene* 16:21-26 (1981).

B. Cloning methods for the isolation of nucleotide sequences encoding T2R family members

In general, the nucleic acid sequences encoding T2R family members and related nucleic acid sequence homologs are cloned from cDNA and genomic DNA libraries by hybridization with probes, or isolated using amplification techniques with oligonucleotide primers. For example, T2R sequences are typically isolated from mammalian nucleic acid (genomic or cDNA) libraries by hybridizing with a nucleic acid probe, the sequence of which can be derived from SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:20, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:34, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:41, SEQ ID NO:43, SEQ ID NO:45, SEQ ID NO:52, SEQ ID NO:54, SEQ ID NO:57, SEQ ID NO:61, SEQ ID NO:63, SEQ ID NO:78, SEQ ID NO:80, SEQ ID NO:82, SEQ ID NO:84, SEQ ID NO:86; SEQ ID NO:88, SEQ ID NO:90, SEQ ID

NO:92, SEQ ID NO:94, SEQ ID NO:96, SEQ ID NO:98, SEQ ID NO:100, SEQ ID NO:102, SEQ ID NO:104 SEQ ID NO:106, SEQ ID NO:108, SEQ ID NO:110, SEQ ID NO:112, SEQ ID NO:114, SEQ ID NO:116, SEQ ID NO:118, SEQ ID NO:120, SEQ ID NO:120, SEQ ID NO:122, SEQ ID NO:124, SEQ ID NO:126, SEQ ID NO:128, SEQ ID NO:130, SEQ ID NO:132, SEQ ID NO:134, SEQ ID NO:136, SEQ ID NO:138, SEQ ID NO:140, SEQ ID NO:142, SEQ ID NO:144, SEQ ID NO:146, SEQ ID NO:148, SEQ ID NO:150, SEQ ID NO:152, SEQ ID NO:154, SEQ ID NO:156, SEQ ID NO:157, SEQ ID NO:159, SEQ ID NO:161, SEQ ID NO:163, or SEQ ID NO:165. A suitable tissue from which RNA and cDNA for T2R family members can be isolated is tongue tissue, optionally taste bud tissues or individual taste cells.

Amplification techniques using primers can also be used to amplify and isolate T2R sequences from DNA or RNA. For example, degenerate primers encoding the following amino acid sequences can be used to amplify a sequence of a T2R gene: SEQ ID NOS: 166, 167, 168, 169, 170, or 171 (*see, e.g.*, Dieffenfach & Dveksler, *PCR Primer: A Laboratory Manual* (1995)). These primers can be used, *e.g.*, to amplify either the full length sequence or a probe of one to several hundred nucleotides, which is then used to screen a mammalian library for full-length T2R clones. As described above, such primers can be used to isolate a full length sequence, or a probe which can then be used to isolated a full length sequence, *e.g.*, from a library.

Nucleic acids encoding T2R can also be isolated from expression libraries using antibodies as probes. Such polyclonal or monoclonal antibodies can be raised using the sequence of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5; SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:35, SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:44, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:62, SEQ ID NO:64, SEQ ID NO:65, SEQ ID NO:66, SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, SEQ ID NO:76, SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81, SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID

NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID NO:158, SEQ ID NO:160, SEQ ID NO:162, or SEQ ID NO:164.

Polymorphic variants, alleles, and interspecies homologs that are substantially identical to a T2R family member can be isolated using T2R nucleic acid probes, and oligonucleotides under stringent hybridization conditions, by screening libraries. Alternatively, expression libraries can be used to clone T2R family members and T2R family member polymorphic variants, alleles, and interspecies homologs, by detecting expressed homologs immunologically with antisera or purified antibodies made against a T2R polypeptide, which also recognize and selectively bind to the T2R homolog.

To make a cDNA library, one should choose a source that is rich in T2R mRNA, *e.g.*, tongue tissue, or isolated taste buds. The mRNA is then made into cDNA using reverse transcriptase, ligated into a recombinant vector, and transfected into a recombinant host for propagation, screening and cloning. Methods for making and screening cDNA libraries are well known (*see, e.g.*, Gubler & Hoffman, *Gene* 25:263-269 (1983); Sambrook *et al., supra*; Ausubel *et al., supra*).

For a genomic library, the DNA is extracted from the tissue and either mechanically sheared or enzymatically digested to yield fragments of about 12-20 kb. The fragments are then separated by gradient centrifugation from undesired sizes and are constructed in bacteriophage lambda vectors. These vectors and phage are packaged *in vitro*. Recombinant phage are analyzed by plaque hybridization as described in Benton & Davis, *Science* 196:180-182 (1977). Colony hybridization is carried out as generally described in Grunstein *et al., Proc. Natl. Acad. Sci. USA.*, 72:3961-3965 (1975).

An alternative method of isolating T2R nucleic acid and its homologs combines the use of synthetic oligonucleotide primers and amplification of an RNA or DNA template (*see* U.S. Patents 4,683,195 and 4,683,202; *PCR Protocols: A Guide to Methods and Applications* (Innis *et al.*, eds, 1990)). Methods such as polymerase chain reaction (PCR) and ligase chain reaction (LCR) can be used to amplify nucleic acid sequences of T2R genes directly from mRNA, from cDNA, from genomic libraries or cDNA libraries. Degenerate oligonucleotides can be designed to amplify T2R family

member homologs using the sequences provided herein. Restriction endonuclease sites can be incorporated into the primers. Polymerase chain reaction or other *in vitro* amplification methods may also be useful, for example, to clone nucleic acid sequences that code for proteins to be expressed, to make nucleic acids to use as probes for detecting the presence of T2R-encoding mRNA in physiological samples, for nucleic acid sequencing, or for other purposes. Genes amplified by the PCR reaction can be purified from agarose gels and cloned into an appropriate vector.

Gene expression of T2R family members can also be analyzed by techniques known in the art, *e.g.*, reverse transcription and amplification of mRNA, isolation of total RNA or poly A^+ RNA, northern blotting, dot blotting, *in situ* hybridization, RNase protection, probing DNA microchip arrays, and the like. In one embodiment, high density oligonucleotide analysis technology (*e.g.*, GeneChip™) is used to identify homologs and polymorphic variants of the GPCRs of the invention. In the case where the homologs being identified are linked to a known disease, they can be used with GeneChip™ as a diagnostic tool in detecting the disease in a biological sample, *see, e.g.*, Gunthand *et al., AIDS Res. Hum. Retroviruses* 14: 869-876 (1998); Kozal *et al., Nat. Med.* 2:753-759 (1996); Matson *et al., Anal. Biochem.* 224:110-106 (1995); Lockhart *et al., Nat. Biotechnol.* 14:1675-1680 (1996); Gingeras *et al., Genome Res.* 8:435-448 (1998); Hacia *et al., Nucleic Acids Res.* 26:3865-3866 (1998).

Synthetic oligonucleotides can be used to construct recombinant T2R genes for use as probes or for expression of protein. This method is performed using a series of overlapping oligonucleotides usually 40- 120 bp in length, representing both the sense and nonsense strands of the gene. These DNA fragments are then annealed, ligated and cloned. Alternatively, amplification techniques can be used with precise primers to amplify a specific subsequence of the T2R nucleic acid. The specific subsequence is then ligated into an expression vector.

The nucleic acid encoding a T2R gene is typically cloned into intermediate vectors before transformation into prokaryotic or eukaryotic cells for replication and/or expression. These intermediate vectors are typically prokaryote vectors, *e.g.*, plasmids, or shuttle vectors.

Optionally, nucleic acids encoding chimeric proteins comprising a T2R polypeptide or domains thereof can be made according to standard techniques. For example, a domain such as a ligand binding domain (*e.g.*, an extracellular domain alone,

an extracellular do[redacted]in plus a transmemberane region, or a transm[redacted]brane region alone), an extracellular domain, a transmembrane domain (*e.g.*, one comprising up toseven transmembrane regions and corresponding extracellular and cytosolic loops), the transmembrane domain and a cytoplasmic domain, an active site, a subunit association region, *etc.*, can be covalently linked to a heterologous protein. For example, an extracellular domain can be linked to a heterologous GPCR transmembrane domain, or a heterologous GPCR extracellular domain can be linked to a transmembrane domain. Other heterologous proteins of choice include, *e.g.*, green fluorescent protein, β-gal, glutamate receptor, and the rhodopsin presequence.

C. Expression in prokaryotes and eukaryotes

To obtain high level expression of a cloned gene or nucleic acid, such as those cDNAs encoding a T2R family member, one typically subclones the T2R sequence into an expression vector that contains a strong promoter to direct transcription, a transcription/translation terminator, and if for a nucleic acid encoding a protein, a ribosome binding site for translational initiation. Suitable bacterial promoters are well known in the art and described, *e.g.*, in Sambrook *et al.* and Ausubel *et al.* Bacterial expression systems for expressing the T2R protein are available in, *e.g.*, *E. coli, Bacillus sp.*, and *Salmonella* (Palva *et al.*, *Gene* 22:229-235 (1983); Mosbach *et al.*, *Nature* 302:543-545 (1983). Kits for such expression systems are commercially available. Eukaryotic expression systems for mammalian cells, yeast, and insect cells are well known in the art and are also commercially available. In one embodiment, the eukaryotic expression vector is an adenoviral vector, an adeno-associated vector, or a retroviral vector.

The promoter used to direct expression of a heterologous nucleic acid depends on the particular application. The promoter is optionally positioned about the same distance from the heterologous transcription start site as it is from the transcription start site in its natural setting. As is known in the art, however, some variation in this distance can be accommodated without loss of promoter function.

In addition to the promoter, the expression vector typically contains a transcription unit or expression cassette that contains all the additional elements required for the expression of the T2R-encoding nucleic acid in host cells. A typical expression cassette thus contains a promoter operably linked to the nucleic acid sequence encoding a

T2R and signals required for efficient polyadenylation of the transcript, ribosome binding sites, and translation termination. The nucleic acid sequence encoding a T2R may typically be linked to a cleavable signal peptide sequence to promote secretion of the encoded protein by the transformed cell. Such signal peptides would include, among others, the signal peptides from tissue plasminogen activator, insulin, and neuron growth factor, and juvenile hormone esterase of *Heliothis virescens*. Additional elements of the cassette may include enhancers and, if genomic DNA is used as the structural gene, introns with functional splice donor and acceptor sites.

In addition to a promoter sequence, the expression cassette should also contain a transcription termination region downstream of the structural gene to provide for efficient termination. The termination region may be obtained from the same gene as the promoter sequence or may be obtained from different genes.

The particular expression vector used to transport the genetic information into the cell is not particularly critical. Any of the conventional vectors used for expression in eukaryotic or prokaryotic cells may be used. Standard bacterial expression vectors include plasmids such as pBR322 based plasmids, pSKF, pET23D, and fusion expression systems such as GST and LacZ. Epitope tags can also be added to recombinant proteins to provide convenient methods of isolation, *e.g.*, c-myc.

Expression vectors containing regulatory elements from eukaryotic viruses are typically used in eukaryotic expression vectors, *e.g.*, SV40 vectors, papilloma virus vectors, and vectors derived from Epstein-Barr virus. Other exemplary eukaryotic vectors include pMSG, pAV009/A$^+$, pMTO10/A$^+$, pMAMneo-5, baculovirus pDSVE, and any other vector allowing expression of proteins under the direction of the SV40 early promoter, SV40 later promoter, metallothionein promoter, murine mammary tumor virus promoter, Rous sarcoma virus promoter, polyhedrin promoter, or other promoters shown effective for expression in eukaryotic cells.

Some expression systems have markers that provide gene amplification such as neomycin, hymidine kinase, hygromycin B phosphotransferase, and dihydrofolate reductase. Alternatively, high yield expression systems not involving gene amplification are also suitable, such as using a baculovirus vector in insect cells, with a sequence encoding a T2R family member under the direction of the polyhedrin promoter or other strong baculovirus promoters.

The elements that are typically included in expression vectors also include a replicon that functions in *E. coli*, a gene encoding antibiotic resistance to permit

selection of bacteria that harbor recombinant plasmids, and unique restriction sites in nonessential regions of the plasmid to allow insertion of eukaryotic sequences. The particular antibiotic resistance gene chosen is not critical, any of the many resistance genes known in the art are suitable. The prokaryotic sequences are optionally chosen such that they do not interfere with the replication of the DNA in eukaryotic cells, if necessary.

Standard transfection methods are used to produce bacterial, mammalian, yeast or insect cell lines that express large quantities of a T2R protein, which are then purified using standard techniques (*see, e.g.*, Colley *et al.*, *J. Biol. Chem.* 264:17619-17622 (1989); *Guide to Protein Purification*, in *Methods in Enzymology*, vol. 182 (Deutscher, ed., 1990)). Transformation of eukaryotic and prokaryotic cells are performed according to standard techniques (*see, e.g.*, Morrison, *J. Bact.* 132:349-351 (1977); Clark-Curtiss & Curtiss, *Methods in Enzymology* 101:347-362 (Wu *et al.*, eds, 1983).

Any of the well known procedures for introducing foreign nucleotide sequences into host cells may be used. These include the use of calcium phosphate transfection, polybrene, protoplast fusion, electroporation, liposomes, microinjection, plasma vectors, viral vectors and any of the other well known methods for introducing cloned genomic DNA, cDNA, synthetic DNA or other foreign genetic material into a host cell (*see, e.g.*, Sambrook *et al.*, *supra*). It is only necessary that the particular genetic engineering procedure used be capable of successfully introducing at least one gene into the host cell capable of expressing a T2R gene.

In one preferred embodiment, a polynucleotide encoding a T2R is operably linked to a EF-1α promoter, *e.g.*, using a pEAK10 mammalian expression vector (Edge Biosystems, MD) is used. Such vectors can be introduced into cells, *e.g.*, HEK-293 cells using any standard method, such as transfection using LipofectAMINE (Lifetechnologies).

After the expression vector is introduced into the cells, the transfected cells are cultured under conditions favoring expression of the T2R family member, which is recovered from the culture using standard techniques identified below.

IV. Purification of T2R polypeptides

Either naturally occurring or recombinant T2R polypeptides can be purified for use in functional assays. Optionally, recombinant T2R polypeptides are purified. Naturally occurring T2R polypeptides are purified, *e.g.*, from mammalian tissue such as tongue tissue, and any other source of a T2R homolog. Recombinant T2R polypeptides are purified from any suitable bacterial or eukaryotic expression system, *e.g.*, CHO cells or insect cells.

T2R proteins may be purified to substantial purity by standard techniques, including selective precipitation with such substances as ammonium sulfate; column chromatography, immunopurification methods, and others (*see, e.g.*, Scopes, *Protein Purification: Principles and Practice* (1982); U.S. Patent No. 4,673,641; Ausubel *et al., supra*; and Sambrook *et al., supra*).

A number of procedures can be employed when recombinant T2R family members are being purified. For example, proteins having established molecular adhesion properties can be reversibly fused to the T2R polypeptide. With the appropriate ligand, a T2R can be selectively adsorbed to a purification column and then freed from the column in a relatively pure form. The fused protein is then removed by enzymatic activity. Finally T2R proteins can be purified using immunoaffinity columns.

A. Purification of T2R protein from recombinant cells

Recombinant proteins are expressed by transformed bacteria or eukaryotic cells such as CHO cells or insect cells in large amounts, typically after promoter induction; but expression can be constitutive. Promoter induction with IPTG is a one example of an inducible promoter system. Cells are grown according to standard procedures in the art. Fresh or frozen cells are used for isolation of protein.

Proteins expressed in bacteria may form insoluble aggregates ("inclusion bodies"). Several protocols are suitable for purification of T2R inclusion bodies. For example, purification of inclusion bodies typically involves the extraction, separation and/or purification of inclusion bodies by disruption of bacterial cells, *e.g.*, by incubation in a buffer of 50 mM TRIS/HCL pH 7.5, 50 mM NaCl, 5 mM $MgCl_2$, 1 mM DTT, 0.1 mM ATP, and 1 mM PMSF. The cell suspension can be lysed using 2-3 passages through a French Press, homogenized using a Polytron (Brinkman Instruments) or sonicated on ice. Alternate methods of lysing bacteria are apparent to those of skill in the art (*see, e.g.*, Sambrook *et al., supra*; Ausubel *et al., supra*).

If necessary, the inclusion bodies are solubilized, and the lysed cell suspension is typically centrifuged to remove unwanted insoluble matter. Proteins that formed the inclusion bodies may be renatured by dilution or dialysis with a compatible buffer. Suitable solvents include, but are not limited to urea (from about 4 M to about 8 M), formamide (at least about 80%, volume/volume basis), and guanidine hydrochloride (from about 4 M to about 8 M). Some solvents which are capable of solubilizing aggregate-forming proteins, for example SDS (sodium dodecyl sulfate), 70% formic acid, are inappropriate for use in this procedure due to the possibility of irreversible denaturation of the proteins, accompanied by a lack of immunogenicity and/or activity. Although guanidine hydrochloride and similar agents are denaturants, this denaturation is not irreversible and renaturation may occur upon removal (by dialysis, for example) or dilution of the denaturant, allowing re-formation of immunologically and/or biologically active protein. Other suitable buffers are known to those skilled in the art. T2R polypeptides are separated from other bacterial proteins by standard separation techniques, *e.g.*, with Ni-NTA agarose resin.

Alternatively, it is possible to purify T2R polypeptides from bacteria periplasm. After lysis of the bacteria, when a T2R protein is exported into the periplasm of the bacteria, the periplasmic fraction of the bacteria can be isolated by cold osmotic shock in addition to other methods known to skill in the art. To isolate recombinant proteins from the periplasm, the bacterial cells are centrifuged to form a pellet. The pellet is resuspended in a buffer containing 20% sucrose. To lyse the cells, the bacteria are centrifuged and the pellet is resuspended in ice-cold 5 mM $MgSO_4$ and kept in an ice bath for approximately 10 minutes. The cell suspension is centrifuged and the supernatant decanted and saved. The recombinant proteins present in the supernatant can be separated from the host proteins by standard separation techniques well known to those of skill in the art.

B. Standard protein separation techniques for purifying T2R polypeptides

Solubility fractionation

Often as an initial step, particularly if the protein mixture is complex, an initial salt fractionation can separate many of the unwanted host cell proteins (or proteins derived from the cell culture media) from the recombinant protein of interest. The preferred salt is ammonium sulfate. Ammonium sulfate precipitates proteins by effectively reducing the amount of water in the protein mixture. Proteins then precipitate

on the basis of their solubility. The more hydrophobic a protein is, the more likely it is to precipitate at lower ammonium sulfate concentrations. A typical protocol includes adding saturated ammonium sulfate to a protein solution so that the resultant ammonium sulfate concentration is between 20-30%. This concentration will precipitate the most hydrophobic of proteins. The precipitate is then discarded (unless the protein of interest is hydrophobic) and ammonium sulfate is added to the supernatant to a concentration known to precipitate the protein of interest. The precipitate is then solubilized in buffer and the excess salt removed if necessary, either through dialysis or diafiltration. Other methods that rely on solubility of proteins, such as cold ethanol precipitation, are well known to those of skill in the art and can be used to fractionate complex protein mixtures.

Size differential filtration

The molecular weight of a T2R protein can be used to isolate it from proteins of greater and lesser size using ultrafiltration through membranes of different pore size (for example, Amicon or Millipore membranes). As a first step, the protein mixture is ultrafiltered through a membrane with a pore size that has a lower molecular weight cut-off than the molecular weight of the protein of interest. The retentate of the ultrafiltration is then ultrafiltered against a membrane with a molecular cut off greater than the molecular weight of the protein of interest. The recombinant protein will pass through the membrane into the filtrate. The filtrate can then be chromatographed as described below.

Column chromatography

T2R proteins can also be separated from other proteins on the basis of its size, net surface charge, hydrophobicity, and affinity for ligands. In addition, antibodies raised against proteins can be conjugated to column matrices and the proteins immunopurified. All of these methods are well known in the art. It will be apparent to one of skill that chromatographic techniques can be performed at any scale and using equipment from many different manufacturers (*e.g.*, Pharmacia Biotech).

V. Immunological detection of T2R polypeptides

In addition to the detection of T2R genes and gene expression using nucleic acid hybridization technology, one can also use immunoassays to detect T2R, *e.g.*, to identify taste receptor cells, especially bitter taste receptor cells, and variants of

T2R family members. Immunoassays can be used to qualitatively or quantitatively analyze the T2R. A general overview of the applicable technology can be found in Harlow & Lane, *Antibodies: A Laboratory Manual* (1988).

A. Antibodies to T2R family members

Methods of producing polyclonal and monoclonal antibodies that react specifically with a T2R family member are known to those of skill in the art (*see, e.g.*, Coligan, *Current Protocols in Immunology* (1991); Harlow & Lane, *supra*; Goding, *Monoclonal Antibodies: Principles and Practice* (2d ed. 1986); and Kohler & Milstein, *Nature* 256:495-497 (1975). Such techniques include antibody preparation by selection of antibodies from libraries of recombinant antibodies in phage or similar vectors, as well as preparation of polyclonal and monoclonal antibodies by immunizing rabbits or mice (*see, e.g.*, Huse *et al.*, *Science* 246:1275-1281 (1989); Ward *et al.*, *Nature* 341:544-546 (1989)).

A number of T2R-comprising immunogens may be used to produce antibodies specifically reactive with a T2R family member. For example, a recombinant T2R protein, or an antigenic fragment thereof, is isolated as described herein. Suitable antigenic regions include, *e.g.*, the conserved motifs that are used to identify members of the T2R family, *i.e.*, SEQ ID NOS:166, 167, 168, 169, 170, and 171. Recombinant protein can be expressed in eukaryotic or prokaryotic cells as described above, and purified as generally described above. Recombinant protein is the preferred immunogen for the production of monoclonal or polyclonal antibodies. Alternatively, a synthetic peptide derived from the sequences disclosed herein and conjugated to a carrier protein can be used an immunogen. Naturally occurring protein may also be used either in pure or impure form. The product is then injected into an animal capable of producing antibodies. Either monoclonal or polyclonal antibodies may be generated, for subsequent use in immunoassays to measure the protein.

Methods of production of polyclonal antibodies are known to those of skill in the art. An inbred strain of mice (*e.g.*, BALB/C mice) or rabbits is immunized with the protein using a standard adjuvant, such as Freund's adjuvant, and a standard immunization protocol. The animal's immune response to the immunogen preparation is monitored by taking test bleeds and determining the titer of reactivity to the T2R. When appropriately high titers of antibody to the immunogen are obtained, blood is collected

from the animal and antisera are prepared. Further fractionation of the antisera to enrich for antibodies reactive to the protein can be done if desired (*see* Harlow & Lane, *supra*).

Monoclonal antibodies may be obtained by various techniques familiar to those skilled in the art. Briefly, spleen cells from an animal immunized with a desired antigen are immortalized, commonly by fusion with a myeloma cell (*see* Kohler & Milstein, *Eur. J. Immunol.* 6:511-519 (1976)). Alternative methods of immortalization include transformation with Epstein Barr Virus, oncogenes, or retroviruses, or other methods well known in the art. Colonies arising from single immortalized cells are screened for production of antibodies of the desired specificity and affinity for the antigen, and yield of the monoclonal antibodies produced by such cells may be enhanced by various techniques, including injection into the peritoneal cavity of a vertebrate host. Alternatively, one may isolate DNA sequences which encode a monoclonal antibody or a binding fragment thereof by screening a DNA library from human B cells according to the general protocol outlined by Huse *et al.*, *Science* 246:1275-1281 (1989).

Monoclonal antibodies and polyclonal sera are collected and titered against the immunogen protein in an immunoassay, for example, a solid phase immunoassay with the immunogen immobilized on a solid support. Typically, polyclonal antisera with a titer of 10^4 or greater are selected and tested for their cross reactivity against non-T2R proteins, or even other T2R family members or other related proteins from other organisms, using a competitive binding immunoassay. Specific polyclonal antisera and monoclonal antibodies will usually bind with a K_d of at least about 0.1 mM, more usually at least about 1 μM, optionally at least about 0.1 μM or better, and optionally 0.01 μM or better.

Once T2R family member specific antibodies are available, individual T2R proteins can be detected by a variety of immunoassay methods. For a review of immunological and immunoassay procedures, see *Basic and Clinical Immunology* (Stites & Terr eds., 7th ed. 1991). Moreover, the immunoassays of the present invention can be performed in any of several configurations, which are reviewed extensively in *Enzyme Immunoassay* (Maggio, ed., 1980); and Harlow & Lane, *supra*.

B. Immunological binding assays

T2R proteins can be detected and/or quantified using any of a number of well recognized immunological binding assays (*see, e.g.*, U.S. Patents 4,366,241; 4,376,110; 4,517,288; and 4,837,168). For a review of the general immunoassays, see

also *Methods in Cell Biology: Antibodies in Cell Biology*, volume [redacted] (Asai, ed. 1993); *Basic and Clinical Immunology* (Stites & Terr, eds., 7th ed. 1991). Immunological binding assays (or immunoassays) typically use an antibody that specifically binds to a protein or antigen of choice (in this case a T2R family member or an antigenic subsequence thereof). The antibody (*e.g.*, anti-T2R) may be produced by any of a number of means well known to those of skill in the art and as described above.

Immunoassays also often use a labeling agent to specifically bind to and label the complex formed by the antibody and antigen. The labeling agent may itself be one of the moieties comprising the antibody/antigen complex. Thus, the labeling agent may be a labeled T2R polypeptide or a labeled anti-T2R antibody. Alternatively, the labeling agent may be a third moiety, such a secondary antibody, that specifically binds to the antibody/T2R complex (a secondary antibody is typically specific to antibodies of the species from which the first antibody is derived). Other proteins capable of specifically binding immunoglobulin constant regions, such as protein A or protein G may also be used as the label agent. These proteins exhibit a strong non-immunogenic reactivity with immunoglobulin constant regions from a variety of species (*see, e.g.,* Kronval *et al., J. Immunol.* 111:1401-1406 (1973); Akerstrom *et al., J. Immunol.* 135:2589-2542 (1985)). The labeling agent can be modified with a detectable moiety, such as biotin, to which another molecule can specifically bind, such as streptavidin. A variety of detectable moieties are well known to those skilled in the art.

Throughout the assays, incubation and/or washing steps may be required after each combination of reagents. Incubation steps can vary from about 5 seconds to several hours, optionally from about 5 minutes to about 24 hours. However, the incubation time will depend upon the assay format, antigen, volume of solution, concentrations, and the like. Usually, the assays will be carried out at ambient temperature, although they can be conducted over a range of temperatures, such as 10°C to 40°C.

Non-competitive assay formats

Immunoassays for detecting a T2R protein in a sample may be either competitive or noncompetitive. Noncompetitive immunoassays are assays in which the amount of antigen is directly measured. In one preferred "sandwich" assay, for example, the anti-T2R antibodies can be bound directly to a solid substrate on which they are

immobilized. These immobilized antibodies then capture the T2R protein present in the test sample. The T2R protein is thus immobilized is then bound by a labeling agent, such as a second T2R antibody bearing a label. Alternatively, the second antibody may lack a label, but it may, in turn, be bound by a labeled third antibody specific to antibodies of the species from which the second antibody is derived. The second or third antibody is typically modified with a detectable moiety, such as biotin, to which another molecule specifically binds, *e.g.*, streptavidin, to provide a detectable moiety.

Competitive assay formats

In competitive assays, the amount of T2R protein present in the sample is measured indirectly by measuring the amount of a known, added (exogenous) T2R protein displaced (competed away) from an anti-T2R antibody by the unknown T2R protein present in a sample. In one competitive assay, a known amount of T2R protein is added to a sample and the sample is then contacted with an antibody that specifically binds to the T2R. The amount of exogenous T2R protein bound to the antibody is inversely proportional to the concentration of T2R protein present in the sample. In a particularly preferred embodiment, the antibody is immobilized on a solid substrate. The amount of T2R protein bound to the antibody may be determined either by measuring the amount of T2R protein present in a T2R/antibody complex, or alternatively by measuring the amount of remaining uncomplexed protein. The amount of T2R protein may be detected by providing a labeled T2R molecule:

A hapten inhibition assay is another preferred competitive assay. In this assay the known T2R protein is immobilized on a solid substrate. A known amount of anti-T2R antibody is added to the sample, and the sample is then contacted with the immobilized T2R. The amount of anti-T2R antibody bound to the known immobilized T2R protein is inversely proportional to the amount of T2R protein present in the sample. Again, the amount of immobilized antibody may be detected by detecting either the immobilized fraction of antibody or the fraction of the antibody that remains in solution. Detection may be direct where the antibody is labeled or indirect by the subsequent addition of a labeled moiety that specifically binds to the antibody as described above.

Cross-reactivity determinations

Immunoassays in the competitive binding format can also be used for crossreactivity determinations. For example, a protein at least partially encoded by SEQ

ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:20, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:34, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:41, SEQ ID NO:43, SEQ ID NO:45, SEQ ID NO:52, SEQ ID NO:54, SEQ ID NO:57, SEQ ID NO:61, SEQ ID NO:63, SEQ ID NO:78, SEQ ID NO:80, SEQ ID NO:82, SEQ ID NO:84, SEQ ID NO:86; SEQ ID NO:88, SEQ ID NO:90, SEQ ID NO:92, SEQ ID NO:94, SEQ ID NO:96, SEQ ID NO:98, SEQ ID NO:100, SEQ ID NO:102, SEQ ID NO:104 SEQ ID NO:106, SEQ ID NO:108, SEQ ID NO:110, SEQ ID NO:112, SEQ ID NO:114, SEQ ID NO:116, SEQ ID NO:118, SEQ ID NO:120, SEQ ID NO:120, SEQ ID NO:122, SEQ ID NO:124, SEQ ID NO:126, SEQ ID NO:128, SEQ ID NO:130, SEQ ID NO:132, SEQ ID NO:134, SEQ ID NO:136, SEQ ID NO:138, SEQ ID NO:140, SEQ ID NO:142, SEQ ID NO:144, SEQ ID NO:146, SEQ ID NO:148, SEQ ID NO:150, SEQ ID NO:152, SEQ ID NO:154, SEQ ID NO:156, SEQ ID NO:157, SEQ ID NO:159, SEQ ID NO:161, SEQ ID NO:163, or SEQ ID NO:165, can be immobilized to a solid support. Proteins (*e.g.*, T2R proteins and homologs) are added to the assay that compete for binding of the antisera to the immobilized antigen. The ability of the added proteins to compete for binding of the antisera to the immobilized protein is compared to the ability of the T2R polypeptide encoded by SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:20, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:34, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:41, SEQ ID NO:43, SEQ ID NO:45, SEQ ID NO:52, SEQ ID NO:54, SEQ ID NO:57, SEQ ID NO:61, SEQ ID NO:63, SEQ ID NO:78, SEQ ID NO:80, SEQ ID NO:82, SEQ ID NO:84, SEQ ID NO:86; SEQ ID NO:88, SEQ ID NO:90, SEQ ID NO:92, SEQ ID NO:94, SEQ ID NO:96, SEQ ID NO:98, SEQ ID NO:100, SEQ ID NO:102, SEQ ID NO:104 SEQ ID NO:106, SEQ ID NO:108, SEQ ID NO:110, SEQ ID NO:112, SEQ ID NO:114, SEQ ID NO:116, SEQ ID NO:118, SEQ ID NO:120, SEQ ID NO:120, SEQ ID NO:122, SEQ ID NO:124, SEQ ID NO:126, SEQ ID NO:128, SEQ ID NO:130, SEQ ID NO:132, SEQ ID NO:134, SEQ ID NO:136, SEQ ID NO:138, SEQ ID NO:140, SEQ ID NO:142, SEQ ID NO:144, SEQ ID NO:146, SEQ ID NO:148, SEQ ID NO:150, SEQ ID NO:152, SEQ ID NO:154, SEQ ID NO:156, SEQ ID NO:157, SEQ ID NO:159, SEQ ID NO:161, SEQ ID NO:163, or SEQ ID NO:165 to compete with itself. The percent crossreactivity for the above proteins is calculated, using standard calculations. Those antisera with less than

10% crossreactivity with each of the added proteins listed above are selected and pooled. The cross-reacting antibodies are optionally removed from the pooled antisera by immunoabsorption with the added considered proteins, *e.g.*, distantly related homologs. In addition, peptides comprising amino acid sequences representing conserved motifs that are used to identify members of the T2R family can be used in cross-reactivity determinations, *i.e.*, SEQ ID NO:166, SEQ ID NO:167, SEQ ID NO:168; SEQ ID NO:169, SEQ ID NO:170, or SEQ ID NO:171.

The immunoabsorbed and pooled antisera are then used in a competitive binding immunoassay as described above to compare a second protein, thought to be perhaps an allele or polymorphic variant of a T2R family member, to the immunogen protein (*i.e.*, T2R protein encoded by SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:20, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:34, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:41, SEQ ID NO:43, SEQ ID NO:45, SEQ ID NO:52, SEQ ID NO:54, SEQ ID NO:57, SEQ ID NO:61, SEQ ID NO:63, SEQ ID NO:78, SEQ ID NO:80, SEQ ID NO:82, SEQ ID NO:84, SEQ ID NO:86; SEQ ID NO:88, SEQ ID NO:90, SEQ ID NO:92, SEQ ID NO:94, SEQ ID NO:96, SEQ ID NO:98, SEQ ID NO:100, SEQ ID NO:102, SEQ ID NO:104 SEQ ID NO:106, SEQ ID NO:108, SEQ ID NO:110, SEQ ID NO:112, SEQ ID NO:114, SEQ ID NO:116, SEQ ID NO:118, SEQ ID NO:120, SEQ ID NO:120, SEQ ID NO:122, SEQ ID NO:124, SEQ ID NO:126, SEQ ID NO:128, SEQ ID NO:130, SEQ ID NO:132, SEQ ID NO:134, SEQ ID NO:136, SEQ ID NO:138, SEQ ID NO:140, SEQ ID NO:142, SEQ ID NO:144, SEQ ID NO:146, SEQ ID NO:148, SEQ ID NO:150, SEQ ID NO:152, SEQ ID NO:154, SEQ ID NO:156, SEQ ID NO:157, SEQ ID NO:159, SEQ ID NO:161, SEQ ID NO:163, or SEQ ID NO:165). In order to make this comparison, the two proteins are each assayed at a wide range of concentrations and the amount of each protein required to inhibit 50% of the binding of the antisera to the immobilized protein is determined. If the amount of the second protein required to inhibit 50% of binding is less than 10 times the amount of the protein encoded by SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:20, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:34, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:41, SEQ ID NO:43, SEQ ID NO:45, SEQ ID NO:52, SEQ ID NO:54, SEQ ID NO:57, SEQ ID NO:61, SEQ ID NO:63, SEQ ID

NO:78, SEQ ID NO:80, SEQ ID NO:82, SEQ ID NO:84, SEQ ID NO:86; SEQ ID NO:88, SEQ ID NO:90, SEQ ID NO:92, SEQ ID NO:94, SEQ ID NO:96, SEQ ID NO:98, SEQ ID NO:100, SEQ ID NO:102, SEQ ID NO:104 SEQ ID NO:106, SEQ ID NO:108, SEQ ID NO:110, SEQ ID NO:112, SEQ ID NO:114, SEQ ID NO:116, SEQ ID NO:118, SEQ ID NO:120, SEQ ID NO:120, SEQ ID NO:122, SEQ ID NO:124, SEQ ID NO:126, SEQ ID NO:128, SEQ ID NO:130, SEQ ID NO:132, SEQ ID NO:134, SEQ ID NO:136, SEQ ID NO:138, SEQ ID NO:140, SEQ ID NO:142, SEQ ID NO:144, SEQ ID NO:146, SEQ ID NO:148, SEQ ID NO:150, SEQ ID NO:152, SEQ ID NO:154, SEQ ID NO:156, SEQ ID NO:157, SEQ ID NO:159, SEQ ID NO:161, SEQ ID NO:163, or SEQ ID NO:165 that is required to inhibit 50% of binding, then the second protein is said to specifically bind to the polyclonal antibodies generated to a T2R immunogen.

Antibodies raised against SEQ ID NOs:166-171 can also be used to prepare antibodies that specifically bind only to GPCRs of the T2R family, but not to GPCRs from other families.

Polyclonal antibodies that specifically bind to a particular member of the T2R family, *e.g.*, T2R01, can be make by subtracting out cross-reactive antibodies using other T2R family members. Species-specific polyclonal antibodies can be made in a similar way. For example, antibodies specific to human T2R01 can be made by subtracting out antibodies that are cross-reactive with orthologous sequences, *e.g.*, rat T2R01 or mouse T2R19.

Other assay formats

Western blot (immunoblot) analysis is used to detect and quantify the presence of T2R protein in the sample. The technique generally comprises separating sample proteins by gel electrophoresis on the basis of molecular weight, transferring the separated proteins to a suitable solid support, (such as a nitrocellulose filter, a nylon filter, or derivatized nylon filter), and incubating the sample with the antibodies that specifically bind the T2R protein. The anti-T2R polypeptide antibodies specifically bind to the T2R polypeptide on the solid support. These antibodies may be directly labeled or alternatively may be subsequently detected using labeled antibodies (*e.g.*, labeled sheep anti-mouse antibodies) that specifically bind to the anti-T2R antibodies.

Other assay formats include liposome immunoassays (LIA), which use liposomes designed to bind specific molecules (*e.g.*, antibodies) and release encapsulated

reagents or marke[redacted] The released chemicals are then detected a[redacted]ding to standard techniques (*see* Monroe *et al.*, *Amer. Clin. Prod. Rev.* 5:34-41 (1986)).

Reduction of non-specific binding

One of skill in the art will appreciate that it is often desirable to minimize non-specific binding in immunoassays. Particularly, where the assay involves an antigen or antibody immobilized on a solid substrate it is desirable to minimize the amount of non-specific binding to the substrate. Means of reducing such non-specific binding are well known to those of skill in the art. Typically, this technique involves coating the substrate with a proteinaceous composition. In particular, protein compositions such as bovine serum albumin (BSA), nonfat powdered milk, and gelatin are widely used with powdered milk being most preferred.

Labels

The particular label or detectable group used in the assay is not a critical aspect of the invention, as long as it does not significantly interfere with the specific binding of the antibody used in the assay. The detectable group can be any material having a detectable physical or chemical property. Such detectable labels have been well-developed in the field of immunoassays and, in general, most any label useful in such methods can be applied to the present invention. Thus, a label is any composition detectable by spectroscopic, photochemical, biochemical, immunochemical, electrical, optical or chemical means. Useful labels in the present invention include magnetic beads (*e.g.*, DYNABEADS™), fluorescent dyes (*e.g.*, fluorescein isothiocyanate, Texas red, rhodamine, and the like), radiolabels (*e.g.*, ^{3}H, ^{125}I, ^{35}S, ^{14}C, or ^{32}P), enzymes (*e.g.*, horse radish peroxidase, alkaline phosphatase and others commonly used in an ELISA), and colorimetric labels such as colloidal gold or colored glass or plastic beads (*e.g.*, polystyrene, polypropylene, latex, *etc.*).

The label may be coupled directly or indirectly to the desired component of the assay according to methods well known in the art. As indicated above, a wide variety of labels may be used, with the choice of label depending on sensitivity required, ease of conjugation with the compound, stability requirements, available instrumentation, and disposal provisions.

Non-radioactive labels are often attached by indirect means. Generally, a ligand molecule (*e.g.*, biotin) is covalently bound to the molecule. The ligand then binds

to another molecul[redacted] (*e.g.*, streptavidin) molecule, which is either [redacted]erently detectable or covalently bound to a signal system, such as a detectable enzyme, a fluorescent compound, or a chemiluminescent compound. The ligands and their targets can be used in any suitable combination with antibodies that recognize a T2R protein, or secondary antibodies that recognize anti-T2R.

The molecules can also be conjugated directly to signal generating compounds, *e.g.*, by conjugation with an enzyme or fluorophore. Enzymes of interest as labels will primarily be hydrolases, particularly phosphatases, esterases and glycosidases, or oxidotases, particularly peroxidases. Fluorescent compounds include fluorescein and its derivatives, rhodamine and its derivatives, dansyl, umbelliferone, *etc.* Chemiluminescent compounds include luciferin, and 2,3-dihydrophthalazinediones, *e.g.*, luminol. For a review of various labeling or signal producing systems that may be used, see U.S. Patent No. 4,391,904.

Means of detecting labels are well known to those of skill in the art. Thus, for example, where the label is a radioactive label, means for detection include a scintillation counter or photographic film as in autoradiography. Where the label is a fluorescent label, it may be detected by exciting the fluorochrome with the appropriate wavelength of light and detecting the resulting fluorescence. The fluorescence may be detected visually, by means of photographic film, by the use of electronic detectors such as charge coupled devices (CCDs) or photomultipliers and the like. Similarly, enzymatic labels may be detected by providing the appropriate substrates for the enzyme and detecting the resulting reaction product. Finally simple colorimetric labels may be detected simply by observing the color associated with the label. Thus, in various dipstick assays, conjugated gold often appears pink, while various conjugated beads appear the color of the bead.

Some assay formats do not require the use of labeled components. For instance, agglutination assays can be used to detect the presence of the target antibodies. In this case, antigen-coated particles are agglutinated by samples comprising the target antibodies. In this format, none of the components need be labeled and the presence of the target antibody is detected by simple visual inspection.

VI. Assays for modulators of T2R family members

A. Assays for T2R protein activity

T2R family members and their alleles and polymorphic variants are G-protein coupled receptors that participate in taste transduction, *e.g.*, bitter taste transduction. The activity of T2R polypeptides can be assessed using a variety of *in vitro* and *in vivo* assays to determine functional, chemical, and physical effects, *e.g.*, measuring ligand binding (*e.g.*, radioactive ligand binding), second messengers (*e.g.*, cAMP, cGMP, IP_3, DAG, or Ca^{2+}), ion flux, phosphorylation levels, transcription levels, neurotransmitter levels, and the like. Furthermore, such assays can be used to test for inhibitors and activators of T2R family members. Modulators can also be genetically altered versions of T2R receptors. Such modulators of taste transduction activity are useful for customizing taste, for example to modify the detection of bitter tastes.

The T2R protein of the assay will typically be selected from a polypeptide having a sequence of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5; SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:35, SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:44, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:62, SEQ ID NO:64, SEQ ID NO:65, SEQ ID NO:66, SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, SEQ ID NO:76, SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81, SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID NO:158, SEQ ID NO:160, SEQ ID NO:162, or SEQ ID NO:164 or conservatively modified variant thereof.

Alternatively, the T2R protein of the assay will be derived from a eukaryote and include an amino acid subsequence having amino acid sequence identity to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5; SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:35, SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:44, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:62, SEQ ID NO:64, SEQ ID NO:65, SEQ ID NO:66, SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, SEQ ID NO:76, SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81, SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID NO:158, SEQ ID NO:160, SEQ ID NO:162, or SEQ ID NO:164. Generally, the amino acid sequence identity will be at least 60%, optionally at least 70% to 85%, optionally at least 90-95%. Optionally, the polypeptide of the assays will comprise a domain of a T2R protein, such as an extracellular domain, transmembrane region, transmembrane domain, cytoplasmic domain, ligand binding domain, subunit association domain, active site, and the like. Either the T2R protein or a domain thereof can be covalently linked to a heterologous protein to create a chimeric protein used in the assays described herein.

Modulators of T2R receptor activity are tested using T2R polypeptides as described above, either recombinant or naturally occurring. The protein can be isolated, expressed in a cell, expressed in a membrane derived from a cell, expressed in tissue or in an animal, either recombinant or naturally occurring. For example, tongue slices, dissociated cells from a tongue, transformed cells, or membranes can b used. Modulation is tested using one of the *in vitro* or *in vivo* assays described herein. Taste transduction can also be examined *in vitro* with soluble or solid state reactions, using a full-length

T2R-GPCR or a chimeric molecule such as an extracellular domain or transmembrane region, or combination therof, of a T2R receptor covalently linked to a heterologous signal transduction domain, or a heterologous extracellular domain and/or transmembrane region covalently linked to the transmembrane and/or cytoplasmic domain of a T2R receptor. Furthermore, ligand-binding domains of the protein of interest can be used *in vitro* in soluble or solid state reactions to assay for ligand binding. In numerous embodiements, a chimeric receptor will be made that comprises all or part of a T2R polypeptide as well an additional sequence that facilitates the localization of the T2R to the membrane, such as a rhodopsin, *e.g.*, an N-terminal fragment of a rhodopsin protein.

Ligand binding to a T2R protein, a domain, or chimeric protein can be tested in solution, in a bilayer membrane, attached to a solid phase, in a lipid monolayer, or in vesicles. Binding of a modulator can be tested using, *e.g.*, changes in spectroscopic characteristics (*e.g.*, fluorescence, absorbance, refractive index) hydrodynamic (*e.g.*, shape), chromatographic, or solubility properties.

Receptor-G-protein interactions can also be examined. For example, binding of the G-protein to the receptor or its release from the receptor can be examined. For example, in the absence of GTP, an activator will lead to the formation of a tight complex of a G protein (all three subunits) with the receptor. This complex can be detected in a variety of ways, as noted above. Such an assay can be modified to search for inhibitors, *e.g.*, by adding an activator to the receptor and G protein in the absence of GTP, which form a tight complex, and then screen for inhibitors by looking at dissociation of the receptor-G protein complex. In the presence of GTP, release of the alpha subunit of the G protein from the other two G protein subunits serves as a criterion of activation.

In particularly preferred embodiments, T2R-Gustducin interactions are monitored as a function of T2R receptor activation. As shown in Example IX, mouse T2R5 shows strong cycloheximide-dependent coupling with Gustducin. Such ligand dependent coupling of T2R receptors with Gustducin can be used as a marker to identify modifiers of any member of the T2R family.

An activated or inhibited G-protein will in turn alter the properties of target enzymes, channels, and other effector proteins. The classic examples are the activation of cGMP phosphodiesterase by transducin in the visual system, adenylate cyclase by the stimulatory G-protein, phospholipase C by Gq and other cognate G proteins, and modulation of diverse channels by Gi and other G proteins. Downstream

consequences can also be examined such as generation of diacyl glycerol and IP3 by phospholipase C, and in turn, for calcium mobilization by IP3.

In a preferred embodiment, a T2R polypeptide is expressed in a eukaryotic cell as a chimeric receptor with a heterologous, chaperone sequence that facilitates its maturation and targeting through the secretory pathway. In a preferred embodiment, the heterologous sequence is a rhodopsin sequence, such as an N-terminal fragment of a rhodopsin. Such chimeric T2R receptors can be expressed in any eukaryotic cell, such as HEK-293 cells. Preferably, the cells comprise a functional G protein, *e.g.*, Gα15, that is capable of coupling the chimeric receptor to an intracellular signaling pathway or to a signaling protein such as phospholipase Cβ. Activation of such chimeric receptors in such cells can be detected using any standard method, such as by detecting changes in intracellular calcium by detecting FURA-2 dependent fluorescence in the cell.

Activated GPCR receptors become substrates for kinases that phosphorylate the C-terminal tail of the receptor (and possibly other sites as well). Thus, activators will promote the transfer of ^{32}P from gamma-labeled GTP to the receptor, which can be assayed with a scintillation counter. The phosphorylation of the C-terminal tail will promote the binding of arrestin-like proteins and will interfere with the binding of G-proteins. The kinase/arrestin pathway plays a key role in the desensitization of many GPCR receptors. For example, compounds that modulate the duration a taste receptor stays active would be useful as a means of prolonging a desired taste or cutting off an unpleasant one. For a general review of GPCR signal transduction and methods of assaying signal transduction, *see, e.g., Methods in Enzymology*, vols. 237 and 238 (1994) and volume 96 (1983); Bourne *et al., Nature* 10:349:117-27 (1991); Bourne *et al.*, Nature 348:125-32 (1990); Pitcher *et al., Annu. Rev. Biochem.* 67:653-92 (1998).

Samples or assays that are treated with a potential T2R protein inhibitor or activator are compared to control samples without the test compound, to examine the extent of modulation. Such assays may be carried out in the presence of a bitter tastant that is known to activate the particular receptor, and modulation of the bitter-tastant-dependent activation monitored. Control samples (untreated with activators or inhibitors) are assigned a relative T2R activity value of 100. Inhibition of a T2R protein is achieved when the T2R activity value relative to the control is about 90%, optionally 50%, optionally 25-0%. Activation of a T2R protein is achieved when the T2R activity value relative to the control is 110%, optionally 150%, 200-500%, or 1000-2000%.

Changes in ion flux may be assessed by determining changes in polarization (*i.e.*, electrical potential) of the cell or membrane expressing a T2R protein. One means to determine changes in cellular polarization is by measuring changes in current (thereby measuring changes in polarization) with voltage-clamp and patch-clamp techniques, *e.g.*, the "cell-attached" mode, the "inside-out" mode, and the "whole cell" mode (*see, e.g.*, Ackerman *et al., New Engl. J. Med.* 336:1575-1595 (1997)). Whole cell currents are conveniently determined using the standard methodology (*see, e.g.*, Hamil *et al., PFlugers. Archiv.* 391:85 (1981). Other known assays include: radiolabeled ion flux assays and fluorescence assays using voltage-sensitive dyes (*see, e.g.*, Vestergarrd-Bogind *et al., J. Membrane Biol.* 88:67-75 (1988); Gonzales & Tsien, *Chem. Biol.* 4:269-277 (1997); Daniel *et al., J. Pharmacol. Meth.* 25:185-193 (1991); Holevinsky *et al., J. Membrane Biology* 137:59-70 (1994)). Generally, the compounds to be tested are present in the range from 1 pM to 100 mM.

The effects of the test compounds upon the function of the polypeptides can be measured by examining any of the parameters described above. Any suitable physiological change that affects GPCR activity can be used to assess the influence of a test compound on the polypeptides of this invention. When the functional consequences are determined using intact cells or animals, one can also measure a variety of effects such as transmitter release, hormone release, transcriptional changes to both known and uncharacterized genetic markers (*e.g.*, northern blots), changes in cell metabolism such as cell growth or pH changes, and changes in intracellular second messengers such as Ca^{2+}, IP3, cGMP, or cAMP.

Preferred assays for G-protein coupled receptors include cells that are loaded with ion or voltage sensitive dyes to report receptor activity. Assays for determining activity of such receptors can also use known agonists and antagonists for other G-protein coupled receptors as negative or positive controls to assess activity of tested compounds. In assays for identifying modulatory compounds (*e.g.*, agonists, antagonists), changes in the level of ions in the cytoplasm or membrane voltage will be monitored using an ion sensitive or membrane voltage fluorescent indicator, respectively. Among the ion-sensitive indicators and voltage probes that may be employed are those disclosed in the Molecular Probes 1997 Catalog. For G-protein coupled receptors, promiscuous G-proteins such as Gα15 and Gα16 can be used in the assay of choice

(Wilkie *et al.*, *Proc. Nat'l Acad. Sci. USA* 88:10049-10053 (1991)). Such promiscuous G-proteins allow coupling of a wide range of receptors.

Receptor activation typically initiates subsequent intracellular events, *e.g.*, increases in second messengers such as IP3, which releases intracellular stores of calcium ions. Activation of some G-protein coupled receptors stimulates the formation of inositol triphosphate (IP3) through phospholipase C-mediated hydrolysis of phosphatidylinositol (Berridge & Irvine, *Nature* 312:315-21 (1984)). IP3 in turn stimulates the release of intracellular calcium ion stores. Thus, a change in cytoplasmic calcium ion levels, or a change in second messenger levels such as IP3 can be used to assess G-protein coupled receptor function. Cells expressing such G-protein coupled receptors may exhibit increased cytoplasmic calcium levels as a result of contribution from both intracellular stores and via activation of ion channels, in which case it may be desirable although not necessary to conduct such assays in calcium-free buffer, optionally supplemented with a chelating agent such as EGTA, to distinguish fluorescence response resulting from calcium release from internal stores.

Other assays can involve determining the activity of receptors which, when activated, result in a change in the level of intracellular cyclic nucleotides, *e.g.*, cAMP or cGMP, by activating or inhibiting enzymes such as adenylate cyclase. There are cyclic nucleotide-gated ion channels, *e.g.*, rod photoreceptor cell channels and olfactory neuron channels that are permeable to cations upon activation by binding of cAMP or cGMP (*see, e.g.*, Altenhofen *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 88:9868-9872 (1991) and Dhallan *et al.*, *Nature* 347:184-187 (1990)). In cases where activation of the receptor results in a decrease in cyclic nucleotide levels, it may be preferable to expose the cells to agents that increase intracellular cyclic nucleotide levels, *e.g.*, forskolin, prior to adding a receptor-activating compound to the cells in the assay. Cells for this type of assay can be made by co-transfection of a host cell with DNA encoding a cyclic nucleotide-crated ion channel, GPCR phosphatase and DNA encoding a receptor (*e.g.*, certain glutamate receptors, muscarinic acetylcholine receptors, dopamine receptors, serotonin receptors, and the like), which, when activated, causes a change in cyclic nucleotide levels in the cytoplasm.

In a preferred embodiment, T2R protein activity is measured by expressing a T2R gene in a heterologous cell with a promiscuous G-protein that links the receptor to a phospholipase C signal transduction pathway (*see* Offermanns & Simon, *J. Biol. Chem.* 270:15175-15180 (1995)). Optionally the cell line is HEK-293 (which does not naturally

express T2R genes and the promiscuous G-protein is Gα15 (Offermanns & Simon, *supra*). Modulation of taste transduction is assayed by measuring changes in intracellular Ca^{2+} levels, which change in response to modulation of the T2R signal transduction pathway via administration of a molecule that associates with a T2R protein. Changes in Ca^{2+} levels are optionally measured using fluorescent Ca^{2+} indicator dyes and fluorometric imaging.

In one embodiment, the changes in intracellular cAMP or cGMP can be measured using immunoassays. The method described in Offermanns & Simon, *J. Biol. Chem.* 270:15175-15180 (1995) may be used to determine the level of cAMP. Also, the method described in Felley-Bosco *et al.*, *Am. J. Resp. Cell and Mol. Biol.* 11:159-164 (1994) may be used to determine the level of cGMP. Further, an assay kit for measuring cAMP and/or cGMP is described in U.S. Patent 4,115,538, herein incorporated by reference.

In another embodiment, phosphatidyl inositol (PI) hydrolysis can be analyzed according to U.S. Patent 5,436,128, herein incorporated by reference. Briefly, the assay involves labeling of cells with ^{3}H-myoinositol for 48 or more hrs. The labeled cells are treated with a test compound for one hour. The treated cells are lysed and extracted in chloroform-methanol-water after which the inositol phosphates were separated by ion exchange chromatography and quantified by scintillation counting. Fold stimulation is determined by calculating the ratio of cpm in the presence of agonist to cpm in the presence of buffer control. Likewise, fold inhibition is determined by calculating the ratio of cpm in the presence of antagonist to cpm in the presence of buffer control (which may or may not contain an agonist).

In another embodiment, transcription levels can be measured to assess the effects of a test compound on signal transduction. A host cell containing a T2R protein of interest is contacted with a test compound for a sufficient time to effect any interactions, and then the level of gene expression is measured. The amount of time to effect such interactions may be empirically determined, such as by running a time course and measuring the level of transcription as a function of time. The amount of transcription may be measured by using any method known to those of skill in the art to be suitable. For example, mRNA expression of the protein of interest may be detected using northern blots or their polypeptide products may be identified using immunoassays. Alternatively, transcription based assays using reporter gene may be used as described in U.S. Patent

5,436,128, herein incorporated by reference. The reporter genes could be, *e.g.*, chloramphenicol acetyltransferase, luciferase, β-galactosidase and alkaline phosphatase. Furthermore, the protein of interest can be used as an indirect reporter via attachment to a second reporter such as green fluorescent protein (*see, e.g.*, Mistili & Spector, *Nature Biotechnology* 15:961-964 (1997)).

The amount of transcription is then compared to the amount of transcription in either the same cell in the absence of the test compound, or it may be compared with the amount of transcription in a substantially identical cell that lacks the protein of interest. A substantially identical cell may be derived from the same cells from which the recombinant cell was prepared but which had not been modified by introduction of heterologous DNA. Any difference in the amount of transcription indicates that the test compound has in some manner altered the activity of the protein of interest.

B. Modulators

The compounds tested as modulators of a T2R family member can be any small chemical compound, or a biological entity, such as a protein, sugar, nucleic acid or lipid. Alternatively, modulators can be genetically altered versions of a T2R gene. Typically, test compounds will be small chemical molecules and peptides. Essentially any chemical compound can be used as a potential modulator or ligand in the assays of the invention, although most often compounds can be dissolved in aqueous or organic (especially DMSO-based) solutions are used. The assays are designed to screen large chemical libraries by automating the assay steps and providing compounds from any convenient source to assays, which are typically run in parallel (*e.g.*, in microtiter formats on microtiter plates in robotic assays). It will be appreciated that there are many suppliers of chemical compounds, including Sigma (St. Louis, MO), Aldrich (St. Louis, MO), Sigma-Aldrich (St. Louis, MO), Fluka Chemika-Biochemica Analytika (Buchs, Switzerland) and the like.

In one preferred embodiment, high throughput screening methods involve providing a combinatorial chemical or peptide library containing a large number of potential therapeutic compounds (potential modulator or ligand compounds). Such "combinatorial chemical libraries" or "ligand libraries" are then screened in one or more assays, as described herein, to identify those library members (particular chemical species

or subclasses) that display a desired characteristic activity. The compounds thus identified can serve as conventional "lead compounds" or can themselves be used as potential or actual therapeutics.

A combinatorial chemical library is a collection of diverse chemical compounds generated by either chemical synthesis or biological synthesis, by combining a number of chemical "building blocks" such as reagents. For example, a linear combinatorial chemical library such as a polypeptide library is formed by combining a set of chemical building blocks (amino acids) in every possible way for a given compound length (i.e., the number of amino acids in a polypeptide compound). Millions of chemical compounds can be synthesized through such combinatorial mixing of chemical building blocks.

Preparation and screening of combinatorial chemical libraries is well known to those of skill in the art. Such combinatorial chemical libraries include, but are not limited to, peptide libraries (*see, e.g.,* U.S. Patent 5,010,175, Furka, *Int. J. Pept. Prot. Res.* 37:487-493 (1991) and Houghton *et al.*, *Nature* 354:84-88 (1991)). Other chemistries for generating chemical diversity libraries can also be used. Such chemistries include, but are not limited to: peptoids (*e.g.*, PCT Publication No. WO 91/19735), encoded peptides (*e.g.*, PCT Publication WO 93/20242), random bio-oligomers (*e.g.*, PCT Publication No. WO 92/00091), benzodiazepines (*e.g.*, U.S. Pat. No. 5,288,514), diversomers such as hydantoins, benzodiazepines and dipeptides (Hobbs *et al.*, *Proc. Nat. Acad. Sci. USA* 90:6909-6913 (1993)), vinylogous polypeptides (Hagihara *et al.*, *J. Amer. Chem. Soc.* 114:6568 (1992)), nonpeptidal peptidomimetics with glucose scaffolding (Hirschmann *et al.*, *J. Amer. Chem. Soc.* 114:9217-9218 (1992)), analogous organic syntheses of small compound libraries (Chen *et al.*, *J. Amer. Chem. Soc.* 116:2661 (1994)), oligocarbamates (Cho *et al.*, *Science* 261:1303 (1993)), and/or peptidyl phosphonates (Campbell *et al.*, *J. Org. Chem.* 59:658 (1994)), nucleic acid libraries (*see* Ausubel, Berger and Sambrook, all *supra*), peptide nucleic acid libraries (*see, e.g.,* U.S. Patent 5,539,083), antibody libraries (*see, e.g.,* Vaughn *et al.*, *Nature Biotechnology*, 14(3):309-314 (1996) and PCT/US96/10287), carbohydrate libraries (*see, e.g.,* Liang *et al.*, *Science*, 274:1520-1522 (1996) and U.S. Patent 5,593,853), small organic molecule libraries (*see, e.g.,* benzodiazepines, Baum C&EN, Jan 18, page 33 (1993); isoprenoids, U.S. Patent 5,569,588; thiazolidinones and metathiazanones, U.S. Patent 5,549,974; pyrrolidines, U.S. Patents 5,525,735 and 5,519,134; morpholino compounds, U.S. Patent 5,506,337; benzodiazepines, 5,288,514, and the like).

Devices for the preparation of combinatorial libraries are commercially available (*see, e.g.,* 357 MPS, 390 MPS, Advanced Chem Tech, Louisville KY, Symphony, Rainin, Woburn, MA, 433A Applied Biosystems, Foster City, CA, 9050 Plus, Millipore, Bedford, MA). In addition, numerous combinatorial libraries are themselves commercially available (*see, e.g.,* ComGenex, Princeton, N.J., Tripos, Inc., St. Louis, MO, 3D Pharmaceuticals, Exton, PA, Martek Biosciences, Columbia, MD, etc.).

C. Solid state and soluble high throughput assays

In one embodiment the invention provide soluble assays using molecules such as a domain such as ligand binding domain, an extracellular domain, a transmembrane domain (*e.g.*, one comprising seven transmembrane regions and cytosolic loops), the transmembrane domain and a cytoplasmic domain, an active site, a subunit association region, *etc.*; a domain that is covalently linked to a heterologous protein to create a chimeric molecule; a T2R protein; or a cell or tissue expressing a T2R protein, either naturally occurring or recombinant. In another embodiment, the invention provides solid phase based *in vitro* assays in a high throughput format, where the domain, chimeric molecule, T2R protein, or cell or tissue expressing the T2R is attached to a solid phase substrate.

In the high throughput assays of the invention, it is possible to screen up to several thousand different modulators or ligands in a single day. In particular, each well of a microtiter plate can be used to run a separate assay against a selected potential modulator, or, if concentration or incubation time effects are to be observed, every 5-10 wells can test a single modulator. Thus, a single standard microtiter plate can assay about 100 (*e.g.*, 96) modulators. If 1536 well plates are used, then a single plate can easily assay from about 100- about 1500 different compounds. It is possible to assay several different plates per day; assay screens for up to about 6,000-20,000 different compounds is possible using the integrated systems of the invention. More recently, microfluidic approaches to reagent manipulation have been developed.

The molecule of interest can be bound to the solid state component, directly or indirectly, via covalent or non covalent linkage, *e.g.*, via a tag. The tag can be any of a variety of components. In general, a molecule which binds the tag (a tag binder) is fixed to a solid support, and the tagged molecule of interest (*e.g.*, the taste transduction molecule of interest) is attached to the solid support by interaction of the tag and the tag binder.

A number of tags and tag binders can be used, based upon known molecular interactions well described in the literature. For example, where a tag has a natural binder, for example, biotin, protein A, or protein G, it can be used in conjunction with appropriate tag binders (avidin, streptavidin, neutravidin, the Fc region of an immunoglobulin, *etc.*) Antibodies to molecules with natural binders such as biotin are also widely available and appropriate tag binders; *see,* SIGMA Immunochemicals 1998 catalogue SIGMA, St. Louis MO).

Similarly, any haptenic or antigenic compound can be used in combination with an appropriate antibody to form a tag/tag binder pair. Thousands of specific antibodies are commercially available and many additional antibodies are described in the literature. For example, in one common configuration, the tag is a first antibody and the tag binder is a second antibody which recognizes the first antibody. In addition to antibody-antigen interactions, receptor-ligand interactions are also appropriate as tag and tag-binder pairs. For example, agonists and antagonists of cell membrane receptors (e.g., cell receptor-ligand interactions such as transferrin, c-kit, viral receptor ligands, cytokine receptors, chemokine receptors, interleukin receptors, immunoglobulin receptors and antibodies, the cadherein family, the integrin family, the selectin family, and the like; *see, e.g.,* Pigott & Power, *The Adhesion Molecule Facts Book I* (1993). Similarly, toxins and venoms, viral epitopes, hormones (e.g., opiates, steroids, etc.), intracellular receptors (e.g. which mediate the effects of various small ligands, including steroids, thyroid hormone, retinoids and vitamin D; peptides), drugs, lectins, sugars, nucleic acids (both linear and cyclic polymer configurations), oligosaccharides, proteins, phospholipids and antibodies can all interact with various cell receptors.

Synthetic polymers, such as polyurethanes, polyesters, polycarbonates, polyureas, polyamides, polyethyleneimines, polyarylene sulfides, polysiloxanes, polyimides, and polyacetates can also form an appropriate tag or tag binder. Many other tag/tag binder pairs are also useful in assay systems described herein, as would be apparent to one of skill upon review of this disclosure.

Common linkers such as peptides, polyethers, and the like can also serve as tags, and include polypeptide sequences, such as poly gly sequences of between about 5 and 200 amino acids. Such flexible linkers are known to persons of skill in the art. For example, poly(ethelyne glycol) linkers are available from Shearwater Polymers, Inc. Huntsville, Alabama. These linkers optionally have amide linkages, sulfhydryl linkages, or heterofunctional linkages.

Tag binders are fixed to solid substrates using any variety of methods currently available. Solid substrates are commonly derivatized or functionalized by exposing all or a portion of the substrate to a chemical reagent which fixes a chemical group to the surface which is reactive with a portion of the tag binder. For example, groups which are suitable for attachment to a longer chain portion would include amines, hydroxyl, thiol, and carboxyl groups. Aminoalkylsilanes and hydroxyalkylsilanes can be used to functionalize a variety of surfaces, such as glass surfaces. The construction of such solid phase biopolymer arrays is well described in the literature. *See, e.g.*, Merrifield, *J. Am. Chem. Soc.* 85:2149-2154 (1963) (describing solid phase synthesis of, *e.g.*, peptides); Geysen *et al., J. Immun. Meth.* 102:259-274 (1987) (describing synthesis of solid phase components on pins); Frank & Doring, *Tetrahedron* 44:60316040 (1988) (describing synthesis of various peptide sequences on cellulose disks); Fodor *et al., Science*, 251:767-777 (1991); Sheldon *et al., Clinical Chemistry* 39(4):718-719 (1993); and Kozal *et al., Nature Medicine* 2(7):753759 (1996) (all describing arrays of biopolymers fixed to solid substrates). Non-chemical approaches for fixing tag binders to substrates include other common methods, such as heat, cross-linking by UV radiation, and the like.

D. Computer-based assays

Yet another assay for compounds that modulate T2R protein activity involves computer assisted drug design, in which a computer system is used to generate a three-dimensional structure of a T2R protein based on the structural information encoded by its amino acid sequence. The input amino acid sequence interacts directly and actively with a preestablished algorithm in a computer program to yield secondary, tertiary, and quaternary structural models of the protein. The models of the protein structure are then examined to identify regions of the structure that have the ability to bind, *e.g.*, ligands. These regions are then used to identify ligands that bind to the protein.

The three-dimensional structural model of the protein is generated by entering protein amino acid sequences of at least 10 amino acid residues or corresponding nucleic acid sequences encoding a T2R polypeptide into the computer system. The nucleotide sequence encoding the polypeptide, or the amino acid sequence thereof, can be any of SEQ ID NO:1-165, and conservatively modified versions thereof. The amino acid sequence represents the primary sequence or subsequence of the protein, which encodes the structural information of the protein. At least 10 residues of the amino acid sequence

(or a nucleotide sequence encoding 10 amino acids) are entered in the computer system from computer keyboards, computer readable substrates that include, but are not limited to, electronic storage media (*e.g.*, magnetic diskettes, tapes, cartridges, and chips), optical media (*e.g.*, CD ROM), information distributed by internet sites, and by RAM. The three-dimensional structural model of the protein is then generated by the interaction of the amino acid sequence and the computer system, using software known to those of skill in the art.

The amino acid sequence represents a primary structure that encodes the information necessary to form the secondary, tertiary and quaternary structure of the protein of interest. The software looks at certain parameters encoded by the primary sequence to generate the structural model. These parameters are referred to as "energy terms," and primarily include electrostatic potentials, hydrophobic potentials, solvent accessible surfaces, and hydrogen bonding. Secondary energy terms include van der Waals potentials. Biological molecules form the structures that minimize the energy terms in a cumulative fashion. The computer program is therefore using these terms encoded by the primary structure or amino acid sequence to create the secondary structural model.

The tertiary structure of the protein encoded by the secondary structure is then formed on the basis of the energy terms of the secondary structure. The user at this point can enter additional variables such as whether the protein is membrane bound or soluble, its location in the body, and its cellular location, *e.g.*, cytoplasmic, surface, or nuclear. These variables along with the energy terms of the secondary structure are used to form the model of the tertiary structure. In modeling the tertiary structure, the computer program matches hydrophobic faces of secondary structure with like, and hydrophilic faces of secondary structure with like.

Once the structure has been generated, potential ligand binding regions are identified by the computer system. Three-dimensional structures for potential ligands are generated by entering amino acid or nucleotide sequences or chemical formulas of compounds, as described above. The three-dimensional structure of the potential ligand is then compared to that of the T2R protein to identify ligands that bind to the protein. Binding affinity between the protein and ligands is determined using energy terms to determine which ligands have an enhanced probability of binding to the protein.

Computer systems are also used to screen for mutations, polymorphic variants, alleles and interspecies homologs of T2R genes. Such mutations can be

associated with disease states or genetic traits. As described above, GeneChip™ and related technology can also be used to screen for mutations, polymorphic variants, alleles and interspecies homologs. Once the variants are identified, diagnostic assays can be used to identify patients having such mutated genes. Identification of the mutated T2R genes involves receiving input of a first nucleic acid or amino acid sequence of a T2R gene, *e.g.*, any of SEQ ID NO:1-165, or conservatively modified versions thereof. The sequence is entered into the computer system as described above. The first nucleic acid or amino acid sequence is then compared to a second nucleic acid or amino acid sequence that has substantial identity to the first sequence. The second sequence is entered into the computer system in the manner described above. Once the first and second sequences are compared, nucleotide or amino acid differences between the sequences are identified. Such sequences can represent allelic differences in various T2R genes, and mutations associated with disease states and genetic traits.

IX. Administration and pharmaceutical compositions

Taste modulators can be administered directly to the mammalian subject for modulation of taste, *e.g.*, modulation of bitter taste, *in vivo*. Administration is by any of the routes normally used for introducing a modulator compound into ultimate contact with the tissue to be treated, optionally the tongue or mouth. The taste modulators are administered in any suitable manner, optionally with pharmaceutically acceptable carriers. Suitable methods of administering such modulators are available and well known to those of skill in the art, and, although more than one route can be used to administer a particular composition, a particular route can often provide a more immediate and more effective reaction than another route.

Pharmaceutically acceptable carriers are determined in part by the particular composition being administered, as well as by the particular method used to administer the composition. Accordingly, there is a wide variety of suitable formulations of pharmaceutical compositions of the present invention (*see, e.g., Remington's Pharmaceutical Sciences*, 17th ed. 1985)).

The taste modulators, alone or in combination with other suitable components, can be made into aerosol formulations (*i.e.*, they can be "nebulized") to be administered via inhalation. Aerosol formulations can be placed into pressurized acceptable propellants, such as dichlorodifluoromethane, propane, nitrogen, and the like.

Formulations suitable for administration include aqueous and non-aqueous solutions, isotonic sterile solutions, which can contain antioxidants, buffers, bacteriostats, and solutes that render the formulation isotonic, and aqueous and non-aqueous sterile suspensions that can include suspending agents, solubilizers, thickening agents, stabilizers, and preservatives. In the practice of this invention, compositions can be administered, for example, by orally, topically, intravenously, intraperitoneally, intravesically or intrathecally. Optionally, the compositions are administered orally or nasally. The formulations of compounds can be presented in unit-dose or multi-dose sealed containers, such as ampules and vials. Solutions and suspensions can be prepared from sterile powders, granules, and tablets of the kind previously described. The modulators can also be administered as part a of prepared food or drug.

The dose administered to a patient, in the context of the present invention should be sufficient to effect a beneficial response in the subject over time. The dose will be determined by the efficacy of the particular taste modulators employed and the condition of the subject, as well as the body weight or surface area of the area to be treated. The size of the dose also will be determined by the existence, nature, and extent of any adverse side-effects that accompany the administration of a particular compound or vector in a particular subject.

In determining the effective amount of the modulator to be administered in a physician may evaluate circulating plasma levels of the modulator, modulator toxicities,, and the production of anti-modulator antibodies. In general, the dose equivalent of a modulator is from about 1 ng/kg to 10 mg/kg for a typical subject.

For administration, taste modulators of the present invention can be administered at a rate determined by the LD-50 of the modulator, and the side-effects of the inhibitor at various concentrations, as applied to the mass and overall health of the subject. Administration can be accomplished via single or divided doses.

VIII. Kits

T2R genes and their homologs are useful tools for identifying taste receptor cells, for forensics and paternity determinations, and for examining taste transduction. T2R family member-specific reagents that specifically hybridize to T2R nucleic acids, such as T2R probes and primers, and T2R specific reagents that specifically bind to a T2R protein, *e.g.*, T2R antibodies are used to examine taste cell expression and taste transduction regulation.

Nucleic acid assays for the presence of DNA and RNA for a T2R family member in a sample include numerous techniques are known to those skilled in the art, such as Southern analysis, northern analysis, dot blots, RNase protection, S1 analysis, amplification techniques such as PCR and LCR, and *in situ* hybridization. In *in situ* hybridization, for example, the target nucleic acid is liberated from its cellular surroundings in such as to be available for hybridization within the cell while preserving the cellular morphology for subsequent interpretation and analysis. The following articles provide an overview of the art of *in situ* hybridization: Singer *et al., Biotechniques* 4:230-250 (1986); Haase *et al., Methods in Virology*, vol. VII, pp. 189-226 (1984); and *Nucleic Acid Hybridization: A Practical Approach* (Hames *et al.*, eds. 1987). In addition, a T2R protein can be detected with the various immunoassay techniques described above. The test sample is typically compared to both a positive control (*e.g.*, a sample expressing a recombinant T2R protein) and a negative control.

The present invention also provides for kits for screening for modulators of T2R family members. Such kits can be prepared from readily available materials and reagents. For example, such kits can comprise any one or more of the following materials: T2R nucleic acids or proteins, reaction tubes, and instructions for testing T2R activity. Optionally, the kit contains a biologically active T2R receptor. A wide variety of kits and components can be prepared according to the present invention, depending upon the intended user of the kit and the particular needs of the user.

EXAMPLES

The following examples are provided by way of illustration only and not by way of limitation. Those of skill in the art will readily recognize a variety of noncritical parameters that could be changed or modified to yield essentially similar results.

Example I--Identification of the T2R gene family

Recent genetic linkage studies in humans identified a locus at 5p15 that is associated with the ability to respond to the bitter substance 6-n-propyl-2-thiouracil (PROP; Reed *et al., Am. J. Hum. Genet.* 64:1478-1480 (1999)). To determine whether differences in PROP sensitivity reflected functional differences in a bitter taste receptor, DNA sequence databases were searched for genes encoding candidate transmembrane proteins at this location. Analysis of open reading frames in 450 kb of DNA spanning six

sequenced human genomic BAC clones(*see, e.g.*, accession number AC003015) from this interval identified a novel GPCR (T2R1) at 5p15.2. T2R1 has seven putative transmembrane segments as well as several conserved residues often present in GPCRs (Probst *et al.*, *DNA Cell. Biol.* 11:1-20 (1992)).

Computer searches using T2R1, and reiterated with T2R1-related sequences, revealed 19 additional human receptors (12 full-length and 7 pseudogenes). Full-length hT2Rs were isolated by PCR amplification of genomic DNA. Full-length hT2Rs were used to probe a rat circumvallate cDNA library (Hoon *et al.*, *Cell*, 96:541-551 (1999)) and mouse BAC filter arrays (Genome Systems) at low stringency (50-55 °C wash in 1 X SSC). Southern hybridization experiments were used to identify a non-redundant set of positive BACs and to order overlapping BACs.

These new receptors, referred to as T2Rs (also known as "SF"), define a novel family of GPCRs that are distantly related to V1R vomeronasal receptors and opsins. In contrast to T1Rs, which belong to the superfamily of GPCRs characterized by a large N-terminal domain (Hoon *et al.*, *Cell*, 96:541-551 (1999)), the T2Rs have only a short extracellular N-terminus. Individual members of the T2R-family exhibit 30-70% amino acid identity, and most share highly conserved sequence motifs in the first three and last transmembrane segments, and also in the second cytoplasmic loop. The most divergent regions between T2Rs are the extracellular segments, extending partway into the transmembrane helices. Presumably, the high degree of variability between T2Rs reflects the need to recognize many structurally diverse ligands. Like many other GPCR genes, T2Rs do not contain introns that interrupt coding regions.

Example II--Organization of human T2R genes.

The identified human T2R genes are localized on three chromosomes, and are often organized as head-to-tail arrays. For example, four receptor genes are clustered within a single PAC clone from 7q31 and nine in a BAC clone from 12p13. There may be more human T2Rs in these arrays, as several additional human T2Rs were found within partially sequenced BAC clones that overlap the 9 gene T2R cluster. Within a given array, the similarity of receptors is highly variable, including both relatively related (*e.g.* hT2R13, hT2R14 and hT2R15), and highly divergent receptors (*e.g.* hT2R3 and hT2R4). This type of organization is mirrored in the mouse (see below), and resembles the genomic organization that has been observed for olfactory receptor genes in humans, mice, flies and worms (Rouquier *et al.*, *Nat. Genet.* 18:243-250 (1998)); Sullivan *et al.*,

PNAS 93:884-888 (1996)); Clyne *et al.*, *Neuron* 22:327-388 (1999); Vosshall *et al.*, *Cell* 96:725-736 (1999)); Troemel *et al.*, *Cell* 83:207-218 (1995)).

To obtain estimates of the size of this gene family, various genomic resources were examined. Analysis of the Genome Sequence Survey database (gss) yielded 12 partial T2R sequences. Because this database represents an essentially random sampling of ~14% of the human genome, this number suggests tha there may be ~90 T2R genes in the human genome. Similar searches of the finished (nr) and unfinished high-throughput human genomic sequence databases (htgs) produced 36 full-length and 15 partial T2R sequences. These databases contain ~50% of the genome sequence, also pointing to ~100 T2R genes in the genome. Recognizing that this analysis may be inaccurate due to the quality of the available databases, and the clustered, non-random distribution of T2Rs in the human genome, it is estimated that the T2R family consists of between 80 to 120 members. However, more than 1/3 of the full-length human T2Rs are pseudogenes; thus, the final number of functional human receptors may be significantly smaller (*i.e.*, 40-80). This is similar to what has been observed for human olfactory receptors; where many of the genes appear to be pseudogenes (Rouquier *et al.*, *Nat. Genet.* 18:243-250 (1998)).

Example III--T2R genes are linked to loci involved in bitter taste

The genetics of sweet and bitter tasting has been extensively studied in mice, where a number of loci influencing responses to sweet and bitter tastants have been mapped by behavioral taste-choice assays (Warren and Lewis, *Nature* 227:77-78 (1970)); Fuller, *J. Hered.* 65:33-66 (1974)). The distal end of mouse chromosome 6 contains a cluster of bitter genes that includes *Soa* (for sucrose octaacetate; Capeless *et al.*, *Behav. Genet.* 22:655-663 (1992)), *Rua* (raffinose undecaacetate; Lush, *Genet. Res.* 47:117-123 (1986)), *Cyx* (cycloheximide; Lush and Holland, *Genet. Res.* 52:207-212 (1988)) and *Qui* (quinine; Lush, *Genet. Res.* 44:151-160 (1984)). Recombination studies indicated that these four loci are closely linked to each other, and to *Prp* (salivary proline rich protein; Azen *et al.*, *Trends Genet.* 2:199-200 (1986)). The human 9 gene T2R cluster contains three interspersed *PRP* genes, and maps to an interval that is homologous with the mouse chromosome 6 bitter cluster.

To define the relationship between the mouse chromosome 6 bitter cluster and T2Rs, a large number of mouse T2R genes were isolated and their genomic organization and physical and genetic map locations were determined. By screening

mouse genomic libraries with human T2Rs, 61 BAC-clones containing 28 mouse T2Rs were isolated. The mouse and human receptors display significant amino acid sequence divergence, but share the sequence motifs common to members of this novel family of receptors. Mouse T2Rs were mapped using a mouse/hamster radiation hybrid panel (Research Genetics), and by examining the strain distribution pattern of single nucleotide polymorphisms in a panel of C57BL/6J x DBA/2J recombinant inbred lines (Jackson Laboratory). These studies showed that the mouse genes are clustered at only a few genomic locations. Each genomic interval containing mouse T2Rs is homologous to one containing its closest human counterpart: mT2R8 and hT2R4, mT2R18 and hT2R16, and mT2R19 and hT2R1. Of these 3 sets of potentially orthologous pairs of human/mouse receptors, both the human T2R1 and T2R16 genes map to locations implicated in human bitter perception (Conneally *et al., Hum. Hered.* 26:267-271 (1976); Reed et al., *Am. J. Hum. Genet.* 64:1478-1480 (1999)). The remaining 25 mT2Rs all map to the distal end of chromosome 6, and are represented by 3 BAC contigs spanning at least 400 kb.

Since *Prp* and the bitter-cluster also map to the distal end of mouse chromosome 6, it was determined whether they *localize* within this array of T2Rs. Analysis of a DBA/2 x C57BL/6 recombinant inbred panel revealed that receptors within all 3 BAC-contigs co-segregate with *Prp* and the bitter cluster. Further, the mouse *Prp* gene was isolated (accession number M23236, containing *D6Mit13*) and shown that it lies within the large chromosome 6 T2R cluster. These results demonstrate that T2Rs are intimately linked to loci implicated in bitter perception.

Example IV--T2Rs are expressed in taste receptor cells

The lingual epithelium contains taste buds in three types of papillae: circumvallate papillae at the very back of the tongue, foliate papillae at the posterior lateral edge of the tongue, and fungiform papillae *dispersed* throughout the front half of the tongue surface. Other parts of the oral cavity also have taste buds; these are particularly prominent in the palate epithelium in an area known as the geschmackstreifen and in the epiglottis. To examine the patterns of expression of T2Rs, *in situ* hybridizations were performed using sections of various taste papillae. To ensure that the probes used were expressed in taste tissue, a rat circumvallate cDNA library was screened, leading to the isolation of 14 rat T2Rs cDNAs, each of which is an ortholog of a mouse genomic clone.

To carry out the *in situ* hybridization, tissue was obtained from adult rats and mice. No sex-specific differences of expression patterns were observed, therefore male and female animals were used interchangeably. Fresh frozen sections (16 μm) were attached to silanized slides and prepared for *in situ* hybridization as described previously (Hoon *et al., Cell*, 96:541-551 (1999)). All *in situ* hybridizations were carried out at high stringency (hybridization, 5 X SSC, 50% formamide, 65-72°C; washing, 0.2 X SSC, 72°C). Signals were developed using alkaline-phosphatase conjugated antibodies to digoxigenin and standard chromogenic substrates (Boehringer Mannheim). Where possible, probes contained extensive 3'-non translated sequence to minimize potential cross-hybridization between T2Rs, which was not observed at the stringency used for *in situ* hybridization.

These experiments demonstrated that T2Rs are selectively expressed in subsets of taste receptor cells of the tongue and palate epithelium. Each receptor hybridizes to an average of 2 cells per taste bud per section. Since the sections used in these experiments contain 1/5-1/3 the depth of a taste bud, this reflects a total of 6-10 positive cells/taste bud/probe (or about 15% of the cells in a taste bud). Examination of serial sections demonstrated that all of the taste buds of the circumvallate papilla contain cells that are positive for each of these probes. Thus far, comparable results have been observed with 11 rat T2Rs, and in mouse sections hybridized with 17 different mT2R probes.

Similar studies in foliate, geschmackstreifen and epiglottis taste buds demonstrated that each receptor probe also labels approximately 15% of the cells in every taste bud. In contrast, T2Rs are rarely expressed in fungiform papillae. Examination of hundreds of fungiform taste buds using 11 different T2R probes demonstrated that less than 10% of all fungiform papillae contain T2R-expressing cells. Interestingly, the few fungiform taste buds that do express T2Rs regularly contain multiple positive cells. In fact, the number of positive cells in these papillae is not significantly different from that seen in taste buds from other regions of the oral cavity. Furthermore, fungiform papillae that contain T2R-expressing cells generally appear clustered. This unexpected finding may provide an important clue about the logic of taste coding. It is known that single fibers of the chorda tympani nerve innervate multiple cells in a fungiform taste bud, and that the same fiber often projects to neighboring papillae (Miller, *J. Comp. Neurol.* 158:155-166 (1974)). Perhaps the non-random distribution of T2R-positive taste receptor

cells and taste buds, fungiform papillae reflect a map of connectivity between similar cells.

Northern analysis and *in situ* hybridization demonstrated that T2Rs are not widely expressed outside taste tissue.

Example V--Individual receptor cells express multiple T2R receptors

The above-described results demonstrated that any given T2R is expressed in ~15% of the cells of circumvallate, foliate and palate taste buds. Given that there are over 30 T2Rs in the rodent genome, a taste cell must express more than one receptor. To determine how many receptors are expressed in any cell, and what fraction of taste receptor cells express T2Rs, the number of circumvallate cells labeled with various mixes of 2, 5 or 10 receptors was compared with those labeled with the corresponding individual probes. By counting positive cells in multiple serial sections, it was determined that the number of taste cells labeled with the mixed probes (~20%) was only slightly larger than that labeled by any individual receptor (~15%). Not surprisingly, the signal intensity was significantly enhanced in the mixed probe hybridizations. Similar results were observed in taste buds from other regions of the oral cavity including the fungiform papillae. To directly demonstrate co-expression, double labeling experiments were carried out using a collection of differentially labeled cRNA probes. For double-label fluorescent detection, probes were labeled either with fluorescein or with digoxigenin. An alkaline-phosphatase conjugated anti-fluorescein antibody (Amersham) and a horseradish-peroxidase conjugated anti-digoxigenin antibody were used in combination with fast-red and tyramide fluorogenic substrates (Boehringer Mannheim and New England Nuclear). In these experiments, the majority of cells were found to express multiple receptors.

Example VI--T2R genes are selectively expressed in gustducin-expressing cells

Previous results had shown that T1Rs are expressed in ~30% of taste receptor cells. *In situ* hybridizations with differentially labeled T1R and T2R probes showed that there is no overlap in the expression of these two classes of receptors. Gustducin is also expressed in a large subset of taste receptor cells, but for the most part is not co-expressed with T1Rs (Hoon *et al.*, *Cell*, 96:541-551 (1999)). To determine if T2Rs are expressed in gustducin cells, *in situ* hybridizations were performed using differentially labeled T2Rs and gustducin riboprobes. These experiments demonstrated

that T2Rs are exclusively expressed in gustducin-positive cells of tongue and palate taste buds.

Approximately 1/3 of the gustducin cells in the circumvallate, foliate and palate taste buds did not label with a mix of 10 T2R probes, suggesting that not all gustducin-expressing cells express T2Rs. These cells may express other, perhaps more distantly related receptors, or could be at a different developmental stage. In fungiform taste buds the situation is quite different. Since only 10% of fungiform taste buds contain T2R positive cells, the great majority of gustducin-positive cells in the front of the tongue do not appear to co-express members of the T2R family of receptors. Therefore, there is likely to be an additional set of receptors expressed in the gustducin-positive cells of fungiform papillae.

Example VII--Functional expression of T2Rs

T2Rs were expressed in conjunction with Gα15, a G-protein α-subunit that has been shown to couple a wide range of receptors to phospholipase Cβ (Offermanns and Simon, *J Biol Chem*, 270:15175-80 (1995); Krautwurst *et al.*, *Cell* 95:917-926 (1998)). In this system, receptor activation leads to increases in intracellular calcium [Ca2+]i, which can be monitored at the single cell level using the FURA-2 calcium-indicator dye (Tsien *et al.*, *Cell Calcium* 6:145-157 (1985)). To test and optimize Gα15 coupling, two different GPCRs, a Gαi-coupled μ-opioid receptor (Reisine, *Neuropharm.* 34:463-472 (1995)) and a Gαq-coupled mGluR1 receptor (Masu *et al.*, *Nature* 349:760-765 (1991)), were used. Transfection of these receptors into HEK-293 cell produced robust, agonist-selective, and Gα15-dependent Ca^{2+} responses (Figure 1).

A number of studies have shown that many GPCRs, in particular sensory receptors, require specific "chaperones" for maturation and targeting through the secretory pathway (Baker *et al.*, *Embo J* 13:4886-4895 (1994); Dwyer *et al.*, *Cell* 93:455-466 (1998)). Recently, Krautwurst *et al.*(*Cell* 95:917-926 (1998)) generated chimeric receptors consisting of the first 20 amino acids of rhodopsin and various rodent olfactory receptors. These were targeted to the plasma membrane and functioned as odorant receptors in HEK-293 cells. To determine whether rhodopsin sequences can also help target T2Rs to the plasma membra, rhodopsin-T2R chimeras (rho-T2Rs) were constructed. Expression of these fusion proteins demonstrated that the first 39 amino

acids of bovine rh●psin are very effective in targeting T2Rs to ● plasma membrane of HEK-293 cells (Figure 2). Similar results were obtained with 11 human and 16 rodent T2Rs (see below). To further enhance the level of T2R expression, rho-T2Rs were placed under the control of a strong EF-1α promoter, and introduced as episomal plasmids into modified HEK-293 cells expressing Gα15 (pEAKrapid cells).

A bridge overlap PCR extension technique was used to generate rho-T2R chimeras, which contain the first 39 amino acids of bovine rhodopsin in frame with human and rodent T2R coding sequences (Mehta and Singh, *Biotechniques* 26:1082-1086 (1999). All receptors were cloned into a pEAK10 mammalian expression vector (Edge Biosystems, MD). Modified HEK-293 cells (PEAK^{rapid} cells; Edge BioSystems, MD) were grown and maintained at 37 °C in UltraCulture medium (Bio Whittaker) supplemented with 5% fetal bovine serum, 100 μg/ml Gentamycin sulphate (Fisher), 1 μg/ml Amphotericin B and 2 mM GlutaMax I (Lifetechnologies). For transfection, cells were seeded onto matrigel coated 24-well culture plates or 35 mm recording chambers. After 24 h at 37 °C, cells were washed in OptiMEM medium (Lifetechnologies) and transfected using LipofectAMINE reagent (Lifetechnologies). Transfection efficiencies were estimated by co-transfection of a GFP reporter plasmid, and were typically >70%. Immunofluoresence staining, and activity assays were performed 36-48 h after transfection.

For immunostaining, transfected cells were grown on coated glass coverslips, fixed for 20 min in ice-cold 2% paraformaldehyde, blocked with 1% BSA, and incubated for 4-6 h at 4 °C in blocking buffer containing a 1:1000 dilution of anti-rhodopsin mAb B6-30 (Hargrave, *et al. Exp Eye Res* 42:363-373 (1986)). Chimeric receptor expression was visualized using FITC-coupled donkey anti-mouse secondary antibodies (Jackson Immunochemical).

Two parallel strategies were employed to identify ligands for T2Rs. In one, a random set of human, rat and mouse T2R receptors were selected and individually tested against a collection of 55 bitter and sweet tastants, including (shown with maximum concentrations tested): 5 mM aristolochic acid, 5 mM atropine, 5 mM brucine, 5 mM caffeic acid, 10 mM caffeine, 1 mM chloroquine, 5 mM cycloheximide, 10 mM denatonium benzoate, 5 mM (-) epicatechin, 10 mM L-leucine, 10 mM L-lysine, 10 mM $MgCl_2$, 5 mM naringin, 10 mM nicotine, 2.5 mM papavarine hydrochloride, 3 mM phenyl thiocarbamide, 10 mM 6-n-propyl thiouracil, 1 mM quinacrine, 1 mM quinine

hydrochloride, 800 [redacted]M raffinose undecaacetate, 3 mM salicin, 5 m[redacted] sparteine, 5 mM strychnine nitrate, 3 mM sucrose octaacetate, 2 mM tetraethyl ammonium chloride, 10 mM L-tyrosine, 5 mM yohimbine, 10 mM each of L-glycine, L-alanine, D-tryptophan, L-phenylalanine, L-arginine, sodium saccharin, aspartame, sodium cyclamate, acesulfame K, 150 mM each of sucrose, lactose, maltose, D-glucose, D-fructose, D-galactose, D-sorbitol, 0.1%monellin, 0.1%thaumatin. Additional sweet tastants were 150 μM alitame, 1.8 mM dulcin, 800 μM stevioside, 1.9 mM cyanosusan, 600 μM neohesperidin dihydrochalcone, 10 mM xylitol, 9.7 mM H-Asp-D-Ala-OTMCP, 70 μM N-Dmb-L-Asp-L-Phe-Ome, and 12 μM N-Dmb-L-Asp-D-Val-(S)-α methylbenzylamide. In these assays, functional coupling was assessed based on four criteria: tastant selectivity, temporal specificity, and receptor- and G protein-dependence. The second strategy relied upon data on the genetics of bitter perception in mice to link candidate receptors with specific tastants.

Nearly 30 years ago, it was first reported that various inbred strains of mice differ in their sensitivity to the bitter compound sucrose-octaacetate (Warren and Lewis, *Nature* 227:77-78 (1970)). Subsequently, a number of studies demonstrated that this strain difference was due to allelic variation at a single genetic locus (Soa) (Whitney and Harder, *Behav Genet* 16:559-574 (1986); Capeless *et al.*, *Behav Genet* 22:655-663 (1992)). These findings were extended to additional loci influencing sensitivity to various bitter tastants, including raffinose undecaacetate (Rua), cycloheximide (Cyx), copper glycinate (Glb), and quinine (Qui) (Lush, *Genet. Res.* 44:151-160 (1984); Lush, *Genet. Res.* 47:117-123 (1986), Lush and Holland, (1988)). Genetic mapping experiments showed that the Soa, Rua, Cyx, Qui and Glb loci are clustered at the distal end of chromosome 6 (Lush and Holland, *Genet. Res.* 52:207-212 (1988); Capeless *et al.*, *Behav Genet* 22:655-663 (1992)). In view of the above-described localization of various T2R genes to bitter-associated loci in mice, T2R receptors from this array were constructed as corresponding rho-mT2R chimeras and individually transfected into HEK-293 cells expressing the promiscuous Gα15 protein. After loading the cells with FURA-2, responses to sucrose octaacetate, raffinose undecaacetate, copper glycinate, quinine, and cycloheximide were assayed.

Transfected cells were washed once in Hank's balanced salt solution with 1 mM sodium pyruvate and 10 mM HEPES, pH 7.4 (assay buffer), and loaded with 2 μM FURA-2 AM (Molecular Probes) for 1 h at room temperature. The loading solution was

removed and cells were incubated in 200 µl of assay buffer for 1 h to allow the cleavage of the AM ester. For most experiments, 24-well tissue culture plates containing cells expressing a single rho-T2R were stimulated with 200 µl of a 2x tastant solution (see next section). $[Ca^{2+}]i$ changes were monitored using a Nikon Diaphot 200 microscope equipped with a 10x/0.5 fluor objective with the TILL imaging system (T.I.L.L Photonics GmbH). Acquisition and analysis of the fluorescence images used TILL-Vision software. Generally, $[Ca^{2+}]i$ was measured for 80 - 120 s by sequentially illuminating cells for 200ms at 340nm and 380nm and monitoring the fluorescence emission at 510nm using a cooled CCD camera. The F_{340}/F_{380} ratio was analyzed to measure $[Ca^{2+}]i$.

Kinetics of activation and deactivation were measured using a bath perfusion system. Cells were seeded onto a 150 µl microperfusion chamber, and test solutions were pressure-ejected with a picospritzer apparatus (General Valve, Inc.). Flow-rate was adjusted to ensure complete exchange of the bath solution within 4-5 s. In the case of mT2R5, the entire camera field was measured since >70% of the cells responded to cycloheximide. For mT2R8 and hT2R4, 100 areas of interest in each were averaged for each experiment.

Cells expressing mT2R5 specifically responded to cycloheximide (Figure 3). The response occurred in nearly all transfected cells and was receptor- and Gα15-dependent because cells lacking either of these components did not trigger [Ca2+]i changes, even at 5000-fold higher cycloheximide concentration. As expected for this coupling system, the tastant-induced increase in [Ca2+]i was due to release from internal stores, since analogous results were obtained in nominally zero [Ca2+]out. The activation of mT2R5 by cycloheximide is very selective, as this receptor did not respond to any other tastants, even at concentrations that far exceeded their biologically relevant range of action (Saroli, *Naturwissenschaften* 71:428-9 (1984); Glendinning, *Behav Neurosci* 113:840-854 (1994))(Figure 4a,b). While cycloheximide is only moderately bitter to humans, it is strongly aversive to rodents with a sensitivity threshold of ~0.25 µM (Kusano *et al.*, *Appl. Exptl. Zool.* 6:40-50 (1971); Lush and Holland, *Genet. Res.* 52:207-212 (1988)). In the cell-based assay described herein, the concentration of cycloheximide required to induce half-maximal response of mT2R5 was 0.5 µM, and the threshold was ~0.2 µM (Figure 4c,d). Notably, this dose-response closely matches the sensitivity range of cycloheximide tasting in mice.

To determine the kinetics of the cycloheximide response, rho-mT2R5 transfected cells were placed on a microperfusion chamber and superfused with test solutions under various conditions. The cells showed robust transient responses to micromolar concentrations of cycloheximide that closely follow application of the stimulus (latency <1 s). As expected, when the tastant was removed, [Ca2+]i returned to baseline. A prolonged exposure to cycloheximide (>10 s) resulted in adaptation: a fast increase of [Ca2+]i followed by a gradual, but incomplete decline to the resting level (Figure 4a). Similarly, successive applications of cycloheximide led to significantly reduced responses, indicative of desensitization (Lefkowitz *et al.*, *Cold Spring Harb Symp Quant Biol* 57:127-133 (1992)). This is likely to occur at the level of the receptor, since responses of a control, co-transfected mGluR1 were not altered during the period of cycloheximide desensitization.

To determine whether other T2Rs are also activated by bitter compounds, 11 rhodopsin-tagged human T2R receptors were assayed by individually transfecting them into HEK-293 cells expressing Gα15. Each transfected line was tested against a battery of bitter and sweet tastants, including amino acids, peptides, and other natural and synthetic compounds. These experiments demonstrated that the intensely bitter tastant denatonium induced a significant transient increase in [Ca2+]i in cells transfected with one of the human candidate taste receptors, hT2R4, but not in control untransfected cells (Figure 3), or in cells transfected with other hT2Rs. The denatonium response had a strong dose-dependency with a threshold of ~100 μM. Interestingly, hT2R4 displayed a limited range of promiscuity since it also responded to high concentrations of the bitter tastant 6-n-propyl-2-thiouracil (PROP) (Figure 5).

If the responses of hT2R4 reflect the *in vivo* function of this receptor, it was hypothesized that similarly tuned receptors might be found in other species. The mouse receptor mT2R8 is a likely ortholog of hT2R4: they share ~ 70% identity, while the next closest receptor is only 40% identical; these two genes are contained in homologous genomic intervals. A rho-mT2R8 chimeric receptor was generated and examined for its response to a wide range of tastants. Indeed, mT2R8, like its human counterpart, is activated by denatonium and by high concentrations of PROP (Figures 3 and 5). No other tastants elicited significant responses from cells expressing mT2R8. Because these two receptors share only 70% identity, the similarity in their responses to bitter compounds attests to their role as orthologous bitter taste receptors.

Example VIII--Cycloheximide non-taster mice have mutations in the mT2R5 taste receptor

The demonstration that mT2R5 functions as a high affinity receptor for cycloheximide suggested that the mT2R5 gene might correspond to the Cyx locus. *In situ* hybridization to tissue sections demonstrated that the expression profile of mT2R5 is indistinguishable between taster and non-taster strains (Figure 6). To determine the linkage between mT2R5 and the Cyx locus, polymorphisms in the mT2R5 gene were identified and their distribution in a recombinant inbred panel from a C57BL/6J (non-taster) x DBA/2J (taster) cross was determined. Tight linkage was found between mT2R5 and the Cyx locus. To test the possibility that mutations in the mT2R5 gene were responsible for the Cyx phenotype, the mT2R5 gene was isolated from several additional well-characterized cycloheximide taster (CBA/Ca, BALB/c, C3H/He) and non-taster (129/Sv) strains and their nucleotide sequences determined. Indeed, as would be expected if mT2R5 functions as the cycloheximide receptor in these strains, all the tasters share the same mT2R5 allele as DBA/2J, while the non-tasters share the C57BL/6 allele, which carries missense mutations (Figure 6), including 3 non-conservative amino acid substitutions (T44I, G155D and L294R).

If the mT2R5 C57BL/6 allele is responsible for the taste deficiency of Cyx mutants, its cycloheximide dose-response might recapitulate the sensitivity shift seen in Cyx mutant strains. Two-bottle preference tests have shown that Cyx taster strains avoid cycloheximide with a threshold of 0.25 μM, while non-tasters have a ~ 8-fold decrease in sensitivity (*e.g.* they, are non-tasters at 1 μM, but strongly avoid cycloheximide at 8 μM). A rho-mT2R5 fusion was constructed with the mT2R5 gene from a non-taster strain, and its dose response compared with that of the receptor from taster strains. Remarkably, mT2R5 from the non-taster strains displays a shift in cycloheximide sensitivity (Figure 4d) that resembles the sensitivity of these strains to this bitter tastant. Taken together, these results validate mT2R5 as a cycloheximide receptor, and strongly suggest that mT2R5 corresponds to the Cyx locus.

Example IX--T2Rs couple to gustducin

The above-described demonstration that T2Rs are co-expressed with gustducin suggests that T2Rs activate this G-protein in response to bitter tastants. To

investigate the selectivity of T2R - G-protein coupling, mT2R5 was chosen for study because its activation by cycloheximide recapitulates mouse taste responses. Rho-tagged mT2R5 and gustducin were prepared using a baculovirus expression system. mT2R5-containing membranes were incubated with various purified G-proteins, including gustducin, and measured tastant-induced GTP-γS binding (Hoon *et al.*, *Biochem J* 309:629-636 (1995)). Specifically, infectious Bacmid containing rhodopsin tagged mT2R5 (DBA/2-allele) was produced using the Bac-to-Bac system (Lifetechnologies, MD). Insect larval cells were infected for 60 h with recombinant Bacmid and membranes were prepared as described previously (Ryba and Tirindelli, *J Biol Chem*, 270:6757-6767 (1995)). Peripheral proteins were removed by treatment with 8 M urea and membranes were resuspended in 10 mM HEPES pH7.5, 1 mM EDTA and 1 mM DTT. The expression of rho-mT2R5 was assessed by Western blot using mAb B6-30 and quantitated by comparison with known amounts of rhodopsin. Approximately 300 pmol of rho-mT2R5 could be obtained from 2 x 10^8 infected cells. Gustducin and $G\beta_1\gamma_8$ heterodimers were isolated as described previously (Hoon *et al.*, *Biochem J* 309:629-636 (1995); Ryba and Tirindelli, *J Biol Chem*, 270:6757-6767 (1995)). Receptor-catalyzed exchange of GDP for GTPγS on gustducin and other G-protein α-subunits was measured in the presence of 10 nM rho-mT2R5, 100 μM GDP, and 20 μM $G\beta_1\gamma_8$. All measurements were made at 15-minute time points, and reflect the initial rate of GTPγS binding.

These GTP-γS binding assays revealed exquisite cycloheximide-dependent coupling of mT2R5 to gustducin (Figure 7). In contrast, no coupling was seen with Gαs, Gαi, Gαq or Gαo. No significant GTPγS binding was observed in the absence of receptor, gustducin or βγ-heterodimers. The high selectivity of T2R5 for gustducin, and the exclusive expression of T2Rs in taste receptor cells that contain gustducin, affirm the hypothesis that T2Rs function as gustducin-linked taste receptors.

All publications and patent applications cited in this specification are herein incorporated by reference as if each individual publication or patent application were specifically and individually indicated to be incorporated by reference.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be readily

apparent to one of [redacted]inary skill in the art in light of the teachings [redacted]this invention that certain changes and modifications may be made thereto without departing from the spirit or scope of the appended claims.

5

WHAT IS CLAIMED IS:

1. A method for identifying a compound that modulates taste signaling in taste cells, the method comprising the steps of:

(i) contacting a taste transduction G-protein coupled receptor polypeptide with the compound, the polypeptide comprising greater than 50% amino acid identity to a sequence selected from the group consisting of SEQ ID NO:166, SEQ ID NO:167, SEQ ID NO:168; SEQ ID NO:169, SEQ ID NO:170, and SEQ ID NO:171; and

(ii) determining the functional effect of the compound upon the polypeptide.

2. The method of claim 1, wherein the polypeptide has G-protein coupled receptor activity.

3. The method of claim 1, wherein the functional effect is a chemical effect.

4. The method of claim 1, wherein the functional effect is a physical effect.

5. The method of claim 1, wherein the functional effect is determined by measuring binding of the compound to an extracellular domain or a transmembrane region of the polypeptide.

6. The method of claim 1, wherein the functional effect is determined by measuring binding of radiolabeled GTP to the polypeptide.

7. The method of claim 1, wherein the polypeptide is recombinant.

8. The method of claim 1, wherein the polypeptide is from a rat, a mouse, or a human.

9. The method of claim 1, wherein the polypeptide is expressed in a cell or cell membrane.

10. The method of claim 9, wherein the functional effect is measured by determining changes in the electrical activity of a cell expressing the polypeptide.

11. The method of claim 9, wherein the functional effect is determined by measuring changes in intracellular cAMP, cGMP, IP3, or Ca^{2+}.

12. The method of claim 11, wherein a change in intracellular Ca^{2+} is detected by detecting a change in FURA-2 dependent fluorescence in the cell.

13. The method of claim 9, wherein the cell is a eukaryotic cell.

14. The method of claim 13, wherein the cell is an HEK-293 cell.

15. The method of claim 9, wherein the polypeptide is a fusion protein comprising at least about 20 consecutive N-terminal amino acids of a rhodopsin protein.

16. The method of claim 15, wherein the rhodopsin protein is a bovine rhodopsin.

17. The method of claim 9, wherein the cell comprises Gα15.

18. The method of claim 9, wherein the polypeptide is contacted with the compound in the presence of a bitter tastant, and wherein a difference in the functional effect of the bitter tastant on the cell in the presence of the compound and the functional effect of the bitter tastant on the cell in the absence of the compound indicates that the compound is capable of modulating taste signaling in taste cells.

19. The method of claim 1, wherein the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5; SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:35, SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:44, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:62, SEQ ID NO:64, SEQ ID NO:65, SEQ ID NO:66, SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, SEQ ID NO:76, SEQ ID NO:77, SEQ ID NO:79, SEQ ID

NO:81, SEQ ID N[redacted]3, SEQ ID NO:85, SEQ ID NO:87, SEQ I[redacted]O:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID NO:158, SEQ ID NO:160, SEQ ID NO:162, and SEQ ID NO:164.

20. A method for identifying a compound that modulates taste signaling in taste cells, the method comprising the steps of:

(i) contacting a taste transduction G-protein coupled receptor polypeptide with the compound, the polypeptide comprising greater than 60% amino acid sequence identity to a sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5; SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:35, SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:44, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:62, SEQ ID NO:64, SEQ ID NO:65, SEQ ID NO:66, SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, SEQ ID NO:76, SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81, SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID NO:158, SEQ ID NO:160, SEQ ID NO:162, and SEQ ID NO:164; and

(ii) determining the functional effect of the compound upon the polypeptide.

21. The method of claim 20, wherein the polypeptide has G-protein coupled receptor activity.

22. The method of claim 20, wherein the functional effect is a chemical effect.

23. The method of claim 20, wherein the functional effect is a physical effect.

24. The method of claim 20, wherein the functional effect is determined by measuring binding of the compound to an extracellular domain or a transmembrane region of the polypeptide.

25. The method of claim 20, wherein the functional effect is determined by measuring binding of radiolabeled GTP to the polypeptide.

26. The method of claim 20, wherein the polypeptide is recombinant.

27. The method of claim 20, wherein the polypeptide is from a rat, a mouse, or a human.

28. The method of claim 20, wherein the polypeptide is expressed in a cell or cell membrane.

29. The method of claim 28, wherein the functional effect is measured by determining changes in the electrical activity of a cell expressing the polypeptide.

30. The method of claim 28, wherein the functional effect is determined by measuring changes in intracellular cAMP, cGMP, IP3, or Ca^{2+}.

31. The method of claim 30, wherein a change in intracellular Ca^{2+} is detected by detecting a change in FURA-2 dependent fluorescence in the cell.

32. The method of claim 28, wherein the cell is a eukaryotic cell.

33. The method of claim 32, wherein the cell is an HEK-293 cell.

34. The method of claim 28, wherein the polypeptide is a fusion protein comprising at least about 20 consecutive N-terminal amino acids of a rhodopsin protein.

35. The method of claim 34, wherein the rhodopsin protein is a bovine rhodopsin.

36. The method of claim 28, wherein the cell comprises Gα15.

37. The method of claim 28, wherein the polypeptide is contacted with the compound in the presence of a bitter tastant, and wherein a difference in the functional effect of the bitter tastant on the cell in the presence of the compound and the functional effect of the bitter tastant on the cell in the absence of the compound indicates that the compound is capable of modulating taste signaling in taste cells.

38. The method of claim 20, wherein the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5; SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:35, SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:44, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:62, SEQ ID NO:64, SEQ ID NO:65, SEQ ID NO:66, SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, SEQ ID NO:76, SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81, SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID

NO:151, SEQ ID [redacted]153, SEQ ID NO:155, SEQ ID NO:158, SE[redacted]D NO:160, SEQ ID NO:162, and SEQ ID NO:164.

39. A method for identifying a compound that modulates taste signaling in taste cells, the method comprising the steps of:

(i) contacting a polypeptide comprising an extracellular domain or a transmembrane region of a taste transduction G-protein coupled receptor with the compound, the extracellular domain or transmembrane region comprising greater than 60% amino acid sequence identity to the extracellular domain or transmembrane region of a polypeptide comprising a sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5; SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:35, SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:44, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:62, SEQ ID NO:64, SEQ ID NO:65, SEQ ID NO:66, SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, SEQ ID NO:76, SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81, SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID NO:158, SEQ ID NO:160, SEQ ID NO:162, and SEQ ID NO:164; and

(ii) determining the functional effect of the compound upon the extracellular domain or transmembrane region.

40. The method of claim 39, wherein the polypeptide comprises an extracellular domain or a transmembrane region that is covalently linked to a heterologous polypeptide, forming a chimeric polypeptide.

41. The method of claim 39, wherein the polypeptide has G-protein coupled receptor activity.

42. The method of claim 39, wherein the polypeptide is linked to a solid phase.

43. The method of claim 42, wherein the polypeptide is covalently linked to a solid phase.

44. The method of claim 39, wherein the functional effect is determined by measuring binding of the compound to the extracellular domain or transmembrane region.

45. The method of claim 39, wherein the polypeptide is recombinant.

46. An isolated nucleic acid encoding a taste transduction G-protein coupled recetor, the receptor comprising greater than 50% amino acid sequence identity to a sequence selected from the group consisting of SEQ ID NO:166, SEQ ID NO:167, SEQ ID NO:168, SEQ ID NO:169, SEQ ID NO:170, and SEQ ID NO:171.

47. The nucleic acid of claim 46, wherein the receptor comprises an amino acid sequence selected from the group consisting of SEQ ID NO:166, SEQ ID NO:167, SEQ ID NO:168, SEQ ID NO:169, SEQ ID NO:170, and SEQ ID NO:171.

48. The nucleic acid of claim 46, wherein the nucleic acid encodes a receptor that specifically binds to polyclonal antibodies generated against a polypeptide having a sequence selected from the group consisting of SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81; SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID

NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID NO:158, SEQ ID NO:160, SEQ ID NO:162, and SEQ ID NO:164, but not to polyclonal antibodies generated against a polypeptide having a sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5; SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:35, SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:44, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:62, SEQ ID NO:64, SEQ ID NO:65, SEQ ID NO:66, SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, and SEQ ID NO:76.

49. The nucleic acid of claim 46, wherein the nucleic acid encodes a receptor that has G-protein coupled receptor activity.

50. The nucleic acid of claim 46, wherein the nucleic acid encodes a receptor comprising an amino acid sequence selected from the group consisting of SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81; SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID NO:158, SEQ ID NO:160, SEQ ID NO:162, and SEQ ID NO:164.

51. The nucleic acid of claim 46, wherein the nucleic acid comprises a nucleotide sequence selected from the group consisting of SEQ ID NO:78, SEQ ID NO:80, SEQ ID NO:82, SEQ ID NO:84, SEQ ID NO:86; SEQ ID NO:88, SEQ ID NO:90, SEQ ID NO:92, SEQ ID NO:94, SEQ ID NO:96, SEQ ID NO:98, SEQ ID

NO:100, SEQ ID NO:102, SEQ ID NO:104 SEQ ID NO:106, SEQ ID NO:108, SEQ ID NO:110, SEQ ID NO:112, SEQ ID NO:114, SEQ ID NO:116, SEQ ID NO:118, SEQ ID NO:120, SEQ ID NO:120, SEQ ID NO:122, SEQ ID NO:124, SEQ ID NO:126, SEQ ID NO:128, SEQ ID NO:130, SEQ ID NO:132, SEQ ID NO:134, SEQ ID NO:136, SEQ ID NO:138, SEQ ID NO:140, SEQ ID NO:142, SEQ ID NO:144, SEQ ID NO:146, SEQ ID NO:148, SEQ ID NO:150, SEQ ID NO:152, SEQ ID NO:154, SEQ ID NO:156, SEQ ID NO:157, SEQ ID NO:159, SEQ ID NO:161, SEQ ID NO:163, and SEQ ID NO:165.

52. The nucleic acid of claim 46, wherein the nucleic acid is from a rat or a mouse.

53. The nucleic acid of claim 46, wherein the nucleic acid encodes a chimeric polypeptide comprising an extracellular domain or a transmembrane region linked to a heterologous polypeptide.

54. An expression vector comprising the nucleic acid of claim 46.

55. An isolated cell comprising the vector of claim 54.

56. An isolated nucleic acid encoding a taste transduction G-protein coupled receptor, wherein the nucleic acid is amplified by primers that selectively hybridize under stringent hybridization conditions to the same sequence as degenerate primer sets encoding amino acid sequences selected from the group consisting of:

(1) E(F/A)(I/V/L)(V/L)G(I/V)(L/V)GN(G/T)FI(V/A)LVNC(I/M)DW (SEQ ID NO:166);
(2) (D/G)(F/L)(I/L)L(T/I)(G/A/S)LAISRI(C/G/F)L (SEQ ID NO:167);
(3) NH(L/F)(S/T/N)(L/I/V)W(F/L)(A/T)T(C/S/N)L(S/N/G)(I/V) (SEQ ID NO:168);
(4) FY(F/C)LKIA(N/S)FS(H/N)(P/S)(L/I/V)FL(W/Y)LK (SEQ ID NO:169);
(5) LLI(I/F/V)SLW(K/R)H(S/T)(K/R)(Q/K)(M/I)(Q/K) (SEQ ID NO:170); and
(6) HS(F/L)(I/V)LI(L/M)(G/S/T)N(P/S/N)KL(K/R)(Q/R) (SEQ ID NO:171).

57. The nucleic acid of claim 56, wherein the receptor comprises an amino acid sequence selected from the group consisting of SEQ ID NO:166, SEQ ID NO:167, SEQ ID NO:168, SEQ ID NO:169, SEQ ID NO:170, and SEQ ID NO:171.

58. The nucleic acid of claim 56, wherein the nucleic acid encodes a receptor that specifically binds to polyclonal antibodies generated against a polypeptide having a sequence selected from the group consisting of SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81; SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:87, SEQ ID NO:89, SEQ

ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID NO:158, SEQ ID NO:160, SEQ ID NO:162, and SEQ ID NO:164, but not to polyclonal antibodies generated against a polypeptide having a sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5; SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:35, SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:44, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:62, SEQ ID NO:64, SEQ ID NO:65, SEQ ID NO:66, SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, and SEQ ID NO:76.

59. The nucleic acid of claim 56, wherein the nucleic acid encodes a receptor that has G-protein coupled receptor activity.

60. The nucleic acid of claim 56, wherein the nucleic acid encodes a receptor comprising an amino acid sequence selected from the group consisting of SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81; SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID NO:158, SEQ ID NO:160, SEQ ID NO:162, and SEQ ID NO:164.

61. The nucleic acid of claim 56, wherein the nucleic acid comprises a nucleotide sequence selected from the group consisting of SEQ ID NO:78, SEQ ID NO:80, SEQ ID NO:82, SEQ ID NO:84, SEQ ID NO:86; SEQ ID NO:88, SEQ ID NO:90, SEQ ID NO:92, SEQ ID NO:94, SEQ ID NO:96, SEQ ID NO:98, SEQ ID NO:100, SEQ ID NO:102, SEQ ID NO:104 SEQ ID NO:106, SEQ ID NO:108, SEQ ID NO:110, SEQ ID NO:112, SEQ ID NO:114, SEQ ID NO:116, SEQ ID NO:118, SEQ ID NO:120, SEQ ID NO:120, SEQ ID NO:122, SEQ ID NO:124, SEQ ID NO:126, SEQ ID NO:128, SEQ ID NO:130, SEQ ID NO:132, SEQ ID NO:134, SEQ ID NO:136, SEQ ID NO:138, SEQ ID NO:140, SEQ ID NO:142, SEQ ID NO:144, SEQ ID NO:146, SEQ ID NO:148, SEQ ID NO:150, SEQ ID NO:152, SEQ ID NO:154, SEQ ID NO:156, SEQ ID NO:157, SEQ ID NO:159, SEQ ID NO:161, SEQ ID NO:163, and SEQ ID NO:165.

62. The nucleic acid of claim 56, wherein the nucleic acid is from a rat or a mouse.

63. The nucleic acid of claim 56, wherein the nucleic acid encodes a chimeric polypeptide comprising an extracellular domain or a transmembrane region linked to a heterologous polypeptide.

64. An expression vector comprising the nucleic acid of claim 56.

65. An isolated cell comprising the vector of claim 64.

66. An isolated nucleic acid encoding a taste transduction G-protein coupled receptor, the receptor comprising greater than 60% amino acid sequence identity to a sequence selected from the group consisting of SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81; SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID NO:158, SEQ ID NO:160, SEQ ID NO:162, and SEQ ID NO:164.

67. The nucleic acid of claim 66, wherein the nucleic acid encodes a receptor that specifically binds to polyclonal antibodies generated against a polypeptide having a sequence selected from the group consisting of SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81; SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID NO:158, SEQ ID NO:160, SEQ ID NO:162, and SEQ ID NO:164, but not to polyclonal antibodies generated against a polypeptide having a sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5; SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:35, SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:44, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:62, SEQ ID NO:64, SEQ ID NO:65, SEQ ID NO:66, SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, and SEQ ID NO:76.

68. The nucleic acid of claim 66, wherein the nucleic acid encodes a receptor that has G-protein coupled receptor activity.

69. The nucleic acid of claim 66, wherein the nucleic acid encodes a receptor comprising an amino acid sequence selected from the group consisting of SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81; SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID

NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID NO:158, SEQ ID NO:160, SEQ ID NO:162, and SEQ ID NO:164.

70. The nucleic acid sequence of claim 66, wherein the nucleic acid comprises a nucleotide sequence selected from the group consisting of SEQ ID NO:78, SEQ ID NO:80, SEQ ID NO:82, SEQ ID NO:84, SEQ ID NO:86; SEQ ID NO:88, SEQ ID NO:90, SEQ ID NO:92, SEQ ID NO:94, SEQ ID NO:96, SEQ ID NO:98, SEQ ID NO:100, SEQ ID NO:102, SEQ ID NO:104 SEQ ID NO:106, SEQ ID NO:108, SEQ ID NO:110, SEQ ID NO:112, SEQ ID NO:114, SEQ ID NO:116, SEQ ID NO:118, SEQ ID NO:120, SEQ ID NO:120, SEQ ID NO:122, SEQ ID NO:124, SEQ ID NO:126, SEQ ID NO:128, SEQ ID NO:130, SEQ ID NO:132, SEQ ID NO:134, SEQ ID NO:136, SEQ ID NO:138, SEQ ID NO:140, SEQ ID NO:142, SEQ ID NO:144, SEQ ID NO:146, SEQ ID NO:148, SEQ ID NO:150, SEQ ID NO:152, SEQ ID NO:154, SEQ ID NO:156, SEQ ID NO:157, SEQ ID NO:159, SEQ ID NO:161, SEQ ID NO:163, and SEQ ID NO:165.

71. The nucleic acid of claim 66, wherein the nucleic acid is from a rat or a mouse.

72. The nucleic acid of claim 66, wherein the nucleic acid encodes a chimeric polypeptide comprising an extracellular domain or transmembrane region linked to a heterologous polypeptide.

73. An expression vector comprising the nucleic acid of claim 66.

74. An isolated cell comprising the vector of claim 73.

75. An isolated nucleic acid encoding a taste transduction G-protein coupled receptor, wherein the nucleic acid specifically hybridizes under highly stringent conditions to a nucleic acid having a sequence selected from the group consisting of SEQ ID NO:78, SEQ ID NO:80, SEQ ID NO:82, SEQ ID NO:84, SEQ ID NO:86; SEQ ID NO:88, SEQ ID NO:90, SEQ ID NO:92, SEQ ID NO:94, SEQ ID NO:96, SEQ ID NO:98, SEQ ID NO:100, SEQ ID NO:102, SEQ ID NO:104 SEQ ID NO:106, SEQ ID NO:108, SEQ ID NO:110, SEQ ID NO:112, SEQ ID NO:114, SEQ ID NO:116, SEQ ID NO:118, SEQ ID NO:120, SEQ ID NO:120, SEQ ID NO:122, SEQ ID NO:124, SEQ ID

NO:126, SEQ ID NO:128, SEQ ID NO:130, SEQ ID NO:132, SEQ ID NO:134, SEQ ID NO:136, SEQ ID NO:138, SEQ ID NO:140, SEQ ID NO:142, SEQ ID NO:144, SEQ ID NO:146, SEQ ID NO:148, SEQ ID NO:150, SEQ ID NO:152, SEQ ID NO:154, SEQ ID NO:156, SEQ ID NO:157, SEQ ID NO:159, SEQ ID NO:161, SEQ ID NO:163, and SEQ ID NO:165, but not to a nucleic acid having a sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:20, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:34, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:41, SEQ ID NO:43, SEQ ID NO:45, SEQ ID NO:52, SEQ ID NO:54, SEQ ID NO:57, SEQ ID NO:61, and SEQ ID NO:63.

76. An isolated nucleic acid encoding a taste transduction G-protein coupled receptor, the receptor comprising greater than 60% amino acid sequence identity to a polypeptide having a sequence selected from the group consisting of SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81; SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID NO:158, SEQ ID NO:160, SEQ ID NO:162, and SEQ ID NO:164, wherein the nucleic acid selectively hybridizes under moderately stringent hybridization conditions to a nucleic acid having a nucleotide sequence selected from the group consisting of SEQ ID NO:78, SEQ ID NO:80, SEQ ID NO:82, SEQ ID NO:84, SEQ ID NO:86; SEQ ID NO:88, SEQ ID NO:90, SEQ ID NO:92, SEQ ID NO:94, SEQ ID NO:96, SEQ ID NO:98, SEQ ID NO:100, SEQ ID NO:102, SEQ ID NO:104 SEQ ID NO:106, SEQ ID NO:108, SEQ ID NO:110, SEQ ID NO:112, SEQ ID NO:114, SEQ ID NO:116, SEQ ID NO:118, SEQ ID NO:120, SEQ ID NO:120, SEQ ID NO:122, SEQ ID NO:124, SEQ ID NO:126, SEQ ID NO:128, SEQ ID NO:130, SEQ ID NO:132, SEQ ID NO:134, SEQ ID NO:136, SEQ ID NO:138, SEQ ID NO:140, SEQ ID NO:142, SEQ ID NO:144, SEQ ID NO:146, SEQ ID NO:148, SEQ ID NO:150, SEQ ID NO:152, SEQ ID NO:154, SEQ ID NO:156, SEQ ID

NO:157, SEQ ID [redacted]159, SEQ ID NO:161, SEQ ID NO:163, an[redacted]EQ ID NO:165, but not to a nucleic acid having a sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:20, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:34, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:41, SEQ ID NO:43, SEQ ID NO:45, SEQ ID NO:52, SEQ ID NO:54, SEQ ID NO:57, SEQ ID NO:61, and SEQ ID NO:63.

77. An isolated taste transduction G-protein coupled receptor, the receptor comprising greater than 50% amino acid sequence identity to a sequence selected from the group consisting of SEQ ID NO:166, SEQ ID NO:167, SEQ ID NO:168, SEQ ID NO:169, SEQ ID NO:170, and SEQ ID NO:171.

78. The isolated receptor of claim 77, wherein the receptor comprises an amino acid sequence selected from the group consisting of SEQ ID NO:166, SEQ ID NO:167, SEQ ID NO:168, SEQ ID NO:169, SEQ ID NO:170, and SEQ ID NO:171.

79. The isolated receptor of claim 77, wherein the receptor has G-protein coupled receptor activity.

80. The isolated receptor of claim 77, wherein the polypeptide is covalently linked to a heterologous polypeptide, forming a chimeric polypeptide.

81. The isolated receptor of claim 80, wherein the chimeric polypeptide has G-protein coupled receptor activity.

82. An antibody that selectively binds to the receptor of claim 77.

83. An isolated taste transduction G-protein coupled receptor, the receptor comprising greater than 60% amino acid sequence identity to a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81; SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID NO:127, SEQ ID

NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID NO:158, SEQ ID NO:160, SEQ ID NO:162, and SEQ ID NO:164.

84. The isolated receptor of claim 83, wherein the receptor specifically binds to polyclonal antibodies generated against a polypeptide having a sequence selected from the group consisting of SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81; SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID NO:158, SEQ ID NO:160, SEQ ID NO:162, and SEQ ID NO:164, but not to polyclonal antibodies generated against a polypeptide having a sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5; SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:35, SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:44, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:62, SEQ ID NO:64, SEQ ID NO:65, SEQ ID NO:66, SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, and SEQ ID NO:76.

85. The isolated receptor of claim 83, wherein the receptor has G-protein coupled receptor activity.

86. The isolated receptor of claim 83, wherein the receptor has an amino acid sequence selected from the group consisting of SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81; SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:87, SEQ ID

NO:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID NO:158, SEQ ID NO:160, SEQ ID NO:162, and SEQ ID NO:164.

87. The isolated receptor of claim 83, wherein the receptor is from a rat or a mouse.

88. The isolated receptor of claim 83, wherein the polypeptide is covalently linked to a heterologous polypeptide, forming a chimeric polypeptide.

89. The isolated receptor of claim 88, wherein the chimeric polypeptide has G-protein coupled receptor activity.

90. An antibody that selectively binds to the receptor of claim 83.

91. An expression cassette comprising a polynucleotide sequence that encodes a human taste transduction G protein coupled receptor, operably linked to a heterologous promoter, wherein the receptor comprises an amino acid sequence comprising greater than 60% amino acid sequence identity to a sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5; SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:35, SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:44, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:62, SEQ ID NO:64, SEQ ID NO:65, SEQ ID NO:66, SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, and SEQ ID NO:76.

92. The expression cassette of claim 91, wherein the polynucleotide encodes a receptor comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5; SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:35, SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:44, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:62, SEQ ID NO:64, SEQ ID NO:65, SEQ ID NO:66, SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, and SEQ ID NO:76.

93. An isolated eukaryotic cell comprising the expression cassette of claim 91.

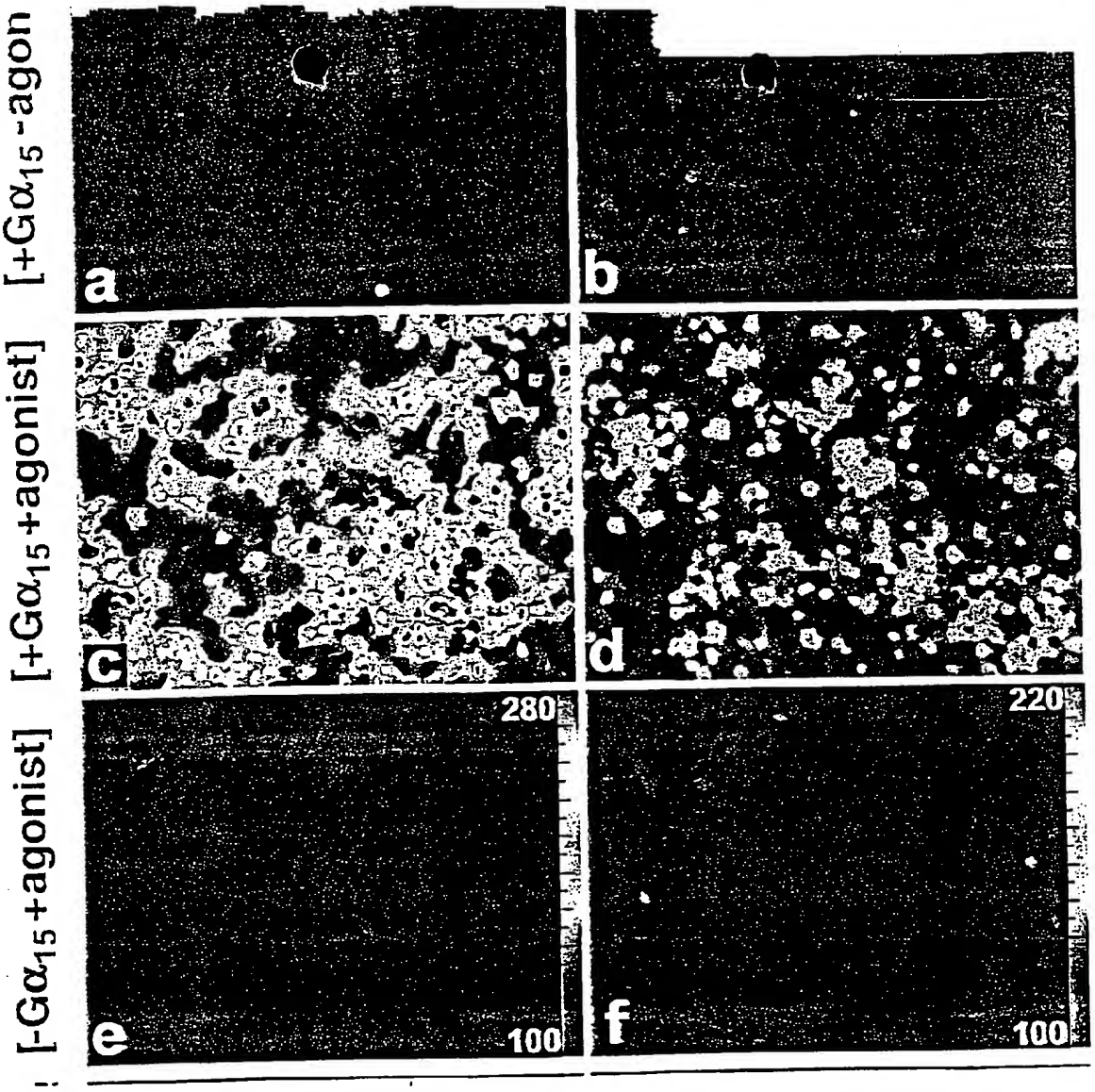


Figure 1

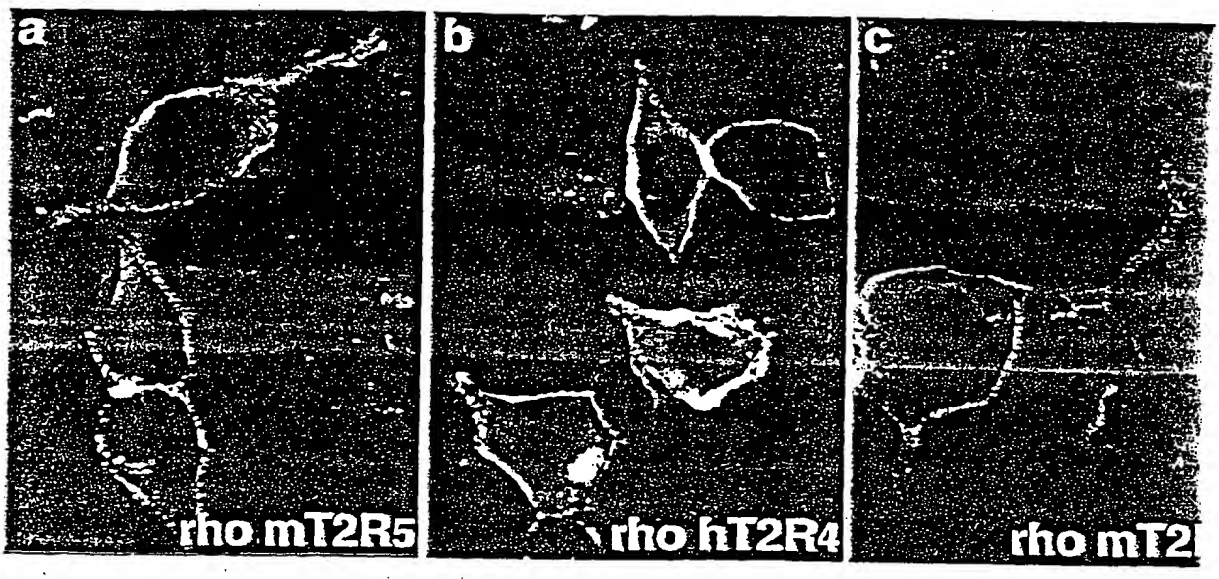


Figure 2

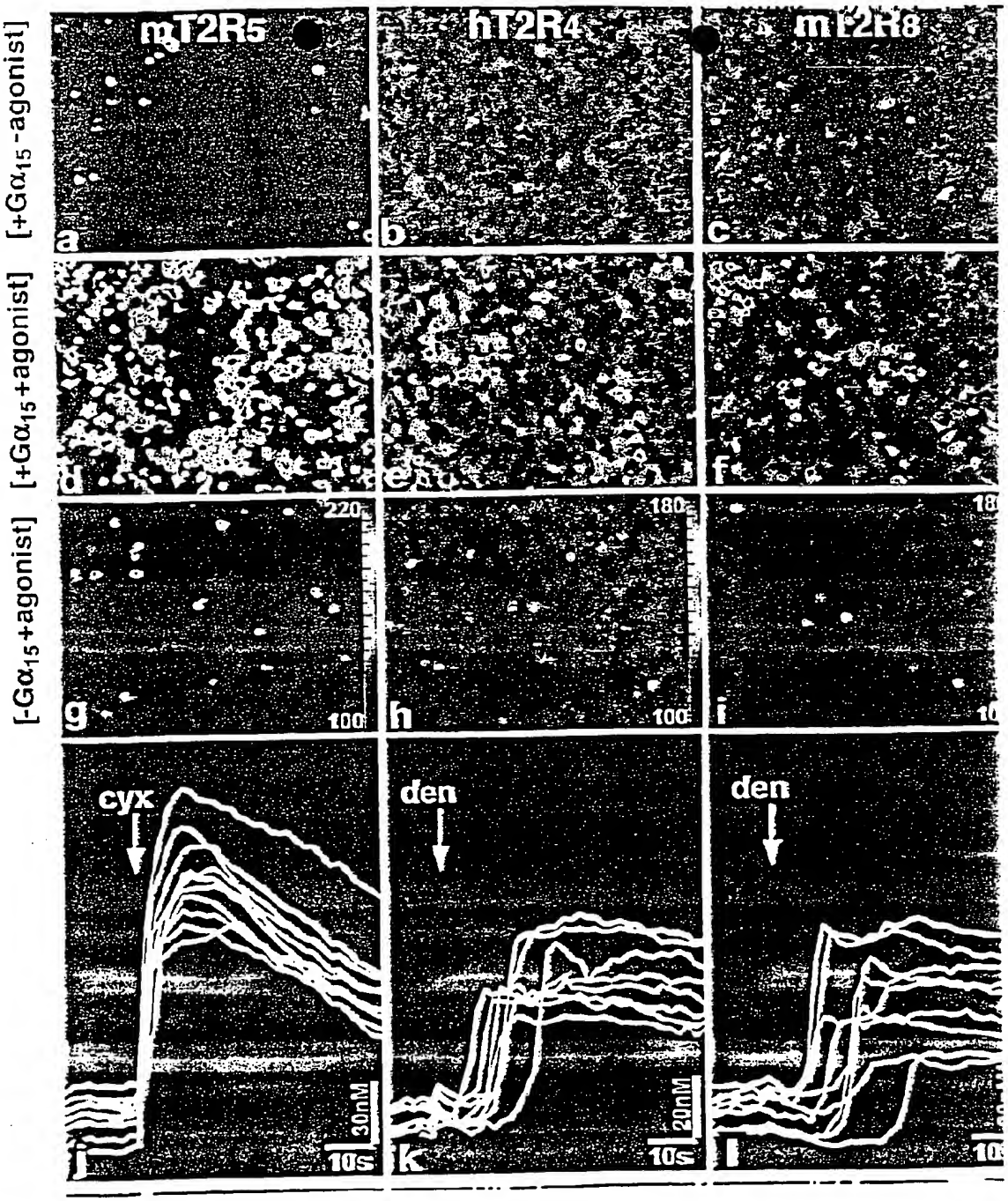
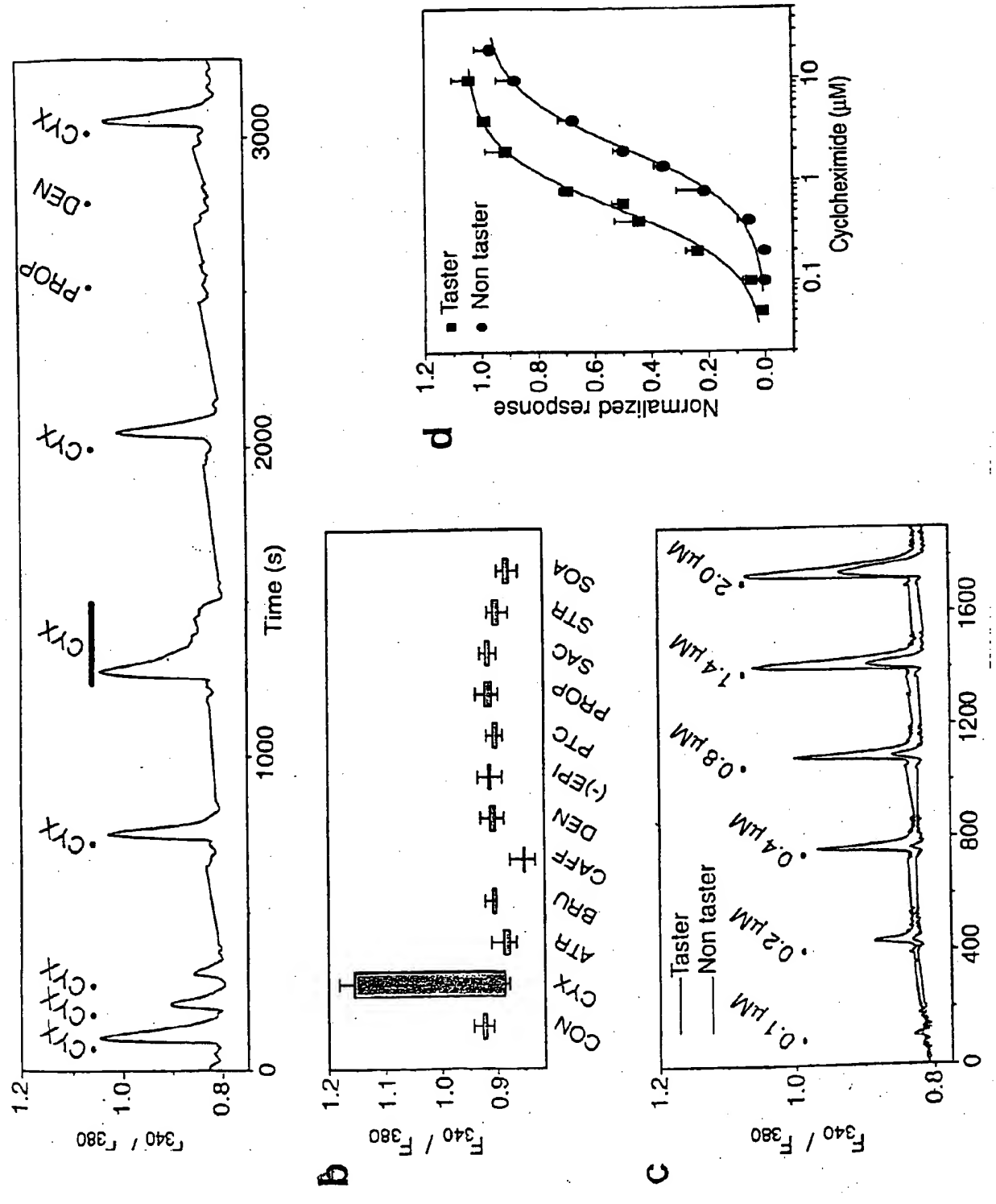
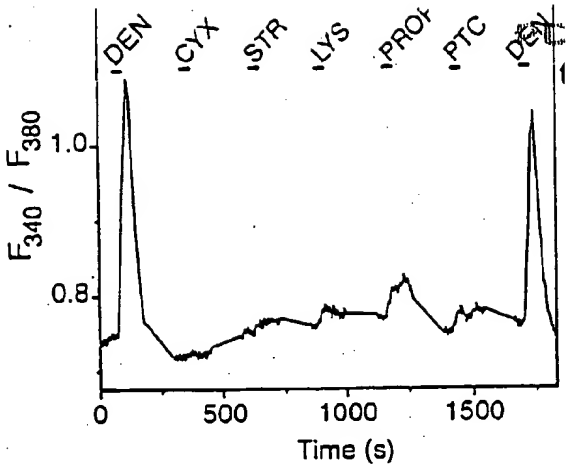
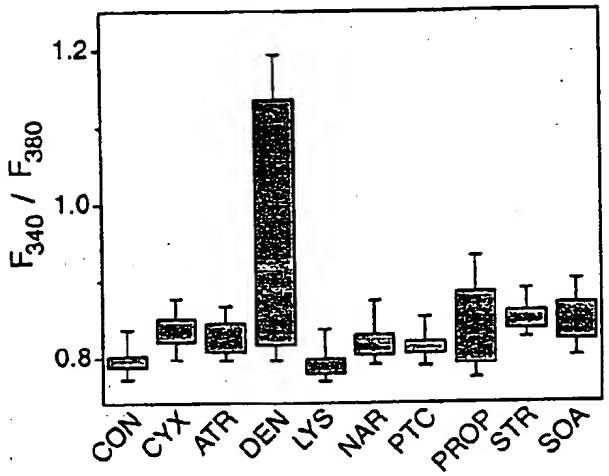

mT2R5
hT2R4
mT2R8
[+Gα15 -agonist]
[+Gα15 +agonist]
[-Gα15 +agonist]
a
b
c
d
e
f
g
h
i
220
180
100
cyx
den
den
30nM
20nM
10s
j
k
l

Figure 3

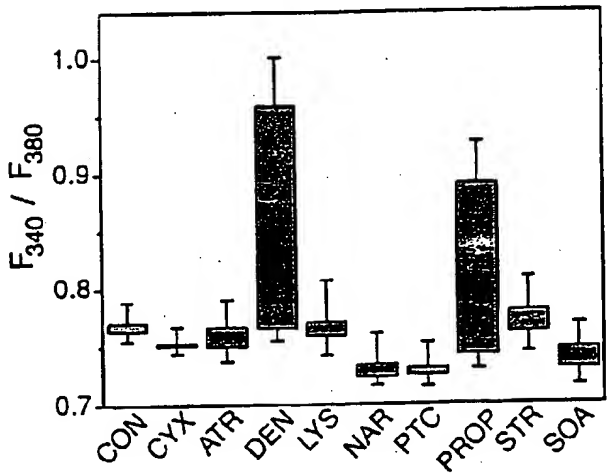

Cyx
Cyx
Cyx
Cyx
Cyx
PROP
DEN
Cyx
Time (s)
0
1000
2000
3000
0.8
1.0
1.2
F340 / F380
a
CON
Cyx
ATR
BRU
CAFF
DEN
(-)EPI
PTC
PROP
SAC
STR
SOA
0.9
1.0
1.1
1.2
F340 / F380
b
0.1 μM
0.2 μM
0.4 μM
0.8 μM
1.4 μM
2.0 μM
Taster
Non taster
0
400
800
1200
1600
0.8
1.0
1.2
F340 / F380
c
Taster
Non taster
0.0
0.2
0.4
0.6
0.8
1.0
1.2
Normalized response
0.1
1
10
Cycloheximide (μM)
d

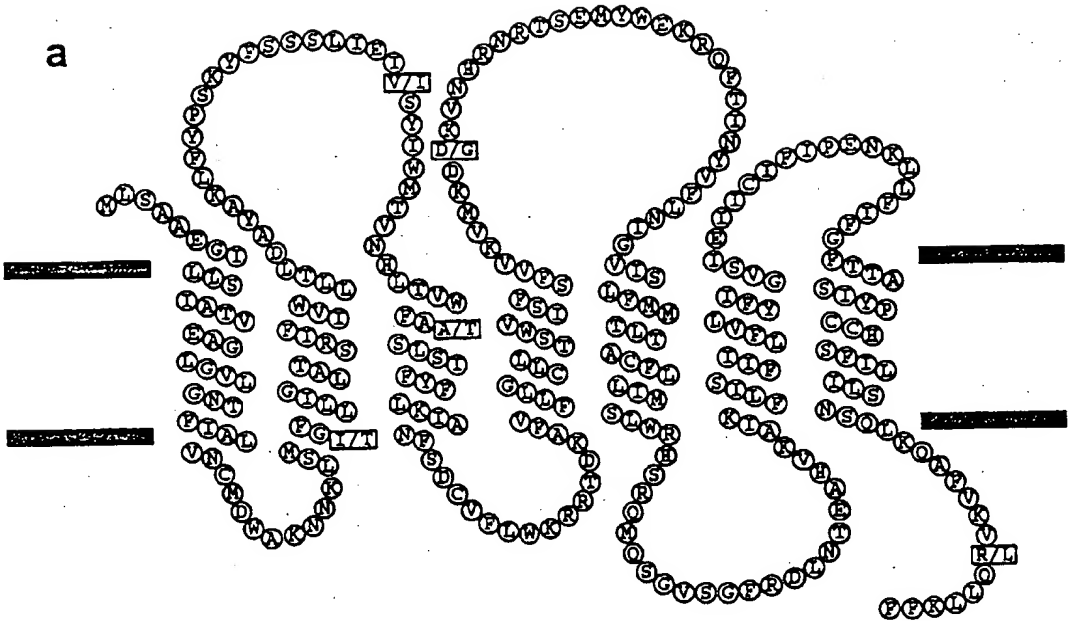

DEN
CYX
STR
LYS
PROP
PTC
DEN
F_{340} / F_{380}
1.0
0.8
0
500
1000
1500
Time (s)

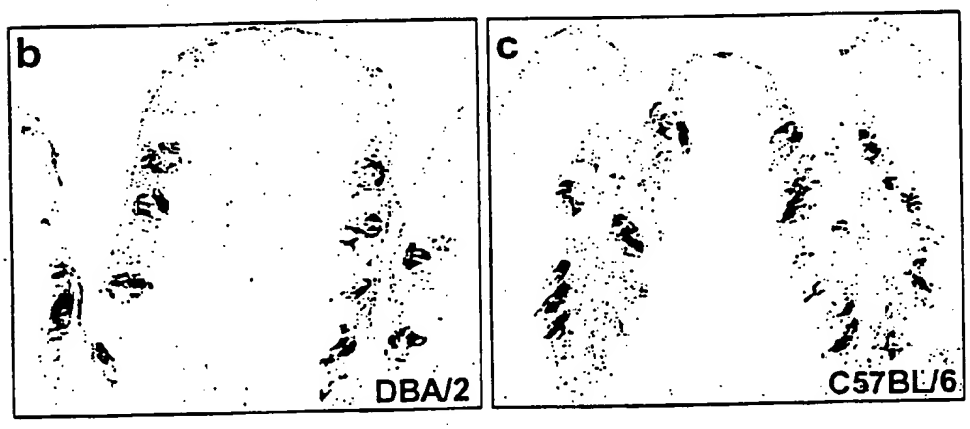
b

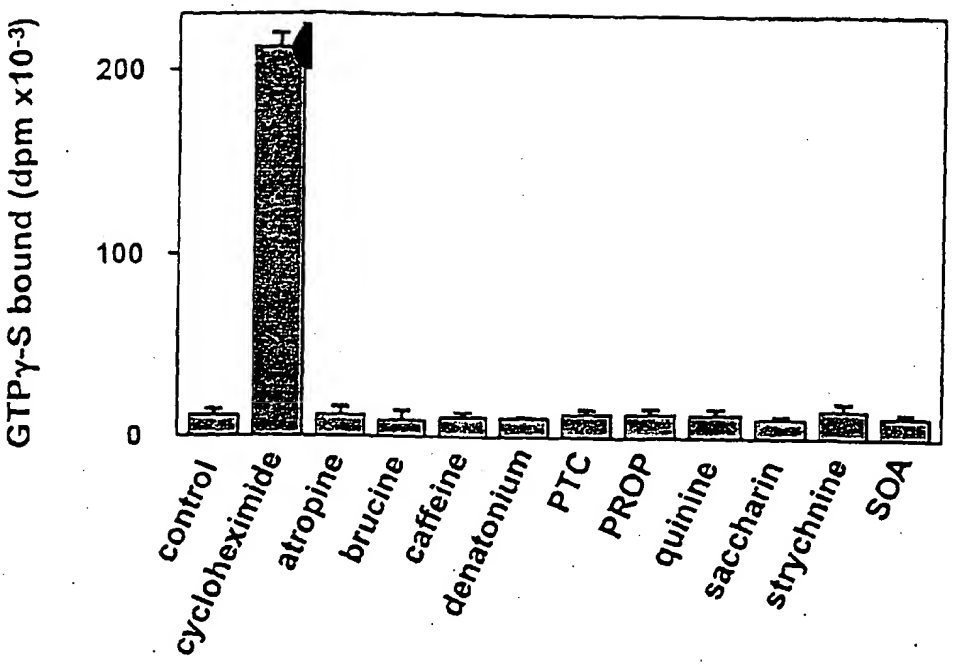

F_{340} / F_{380}
1.2
1.0
0.8
CON
CYX
ATR
DEN
LYS
NAR
PTC
PROP
STR
SOA

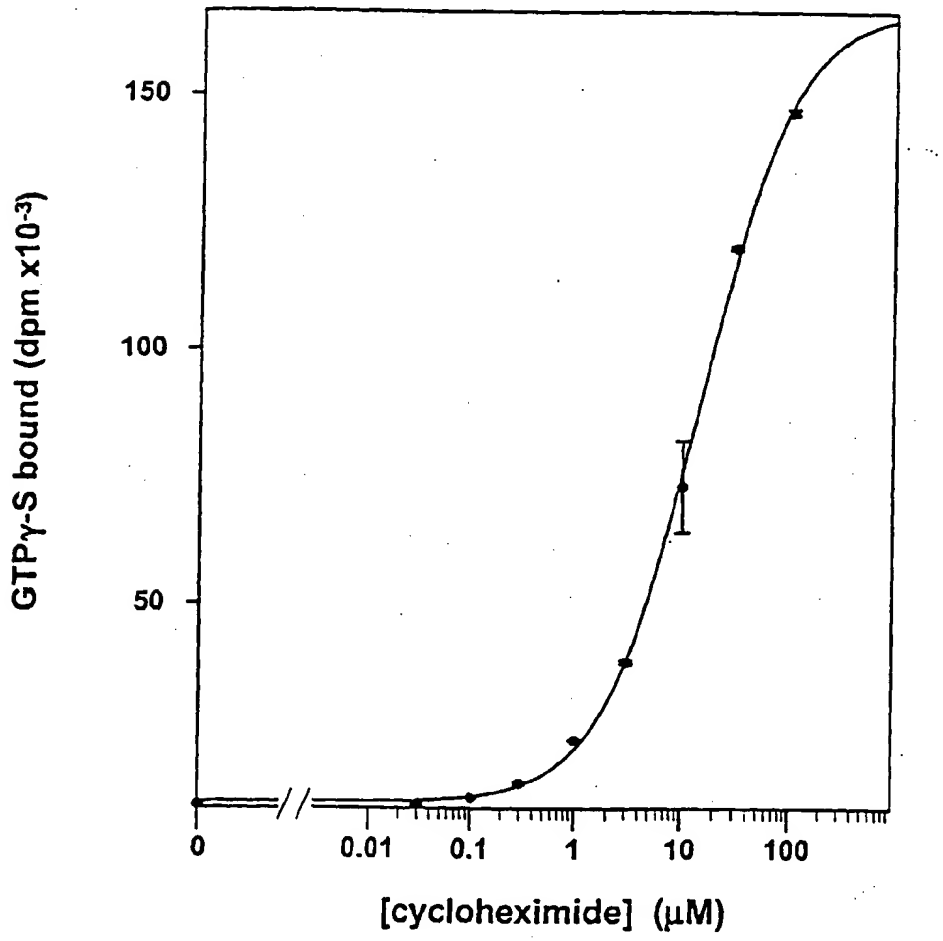
c


F_{340} / F_{380}
1.0
0.9
0.8
0.7
CON
CYX
ATR
DEN
LYS
NAR
PTC
PROP
STR
SOA


a
V/I
D/G
A/T
I/T
R/L


b
DBA/2
c
C57BL/6


GTPγ-S bound (dpm x10-3)
200
100
0
control
cycloheximide
atropine
brucine
caffeine
denatonium
PTC
PROP
quinine
saccharin
strychnine
SOA


GTPγ-S bound (dpm x10-3)
150
100
50
0
0.01
0.1
1
10
100
[cycloheximide] (μM)

T2R ("GR") Family

(hGR=human family members; mGR=mouse family members; rGR=rat family members)
aa=amino acid sequence
nt=nucleotide sequence

>hGR01 aa	>hGR01 nt
MLESHLIIYFLLAVIQFLLGIFTNG IIVVVNGIDLIKHRKMAPLDLLLSC LAVSRIFLQLFIFYVNVIVIFFIEF IMCSANCAILLFINELELWLATWLG VFYCAKVASVRHPLFIWLKMRISKL VPWMILGSLLYVSMICVFHSKYAGF MVPYFLRKFFSQNATIQKEDTLAIQ IFSFVAEFSVPLLIFLFAVLLLIFS LGRHTRQMRNTVAGSRVPGRGAPIS ALLSILSFLILYFSHCMIKVFLSSL KFHIRRFIFLFFILVIGIYPSGHSL ILILGNPKLKQNAKKFLLHSKCCQ	ATGCTAGAGTCTCACCTCATTATCTATTTTCTTCTTGCAGTGATACAATT TCTTCTTGGGATTTTCACAAATGGCATCATTGTGGTGGTGAATGGCATTG ACTTGATCAAGCACAGAAAAATGGCTCCGCTGGATCTCCTTCTTTCTTGT CTGGCAGTTTCTAGAATTTTTCTGCAGTTGTTCATCTTCTACGTTAATGT GATTGTTATCTTCTTCATAGAATTCATCATGTGTTCTGCGAATTGTGCAA TTCTCTTATTTATAAATGAATTGGAACTTTGGCTTGCCACATGGCTCGGC GTTTTCTATTGTGCCAAGGTTGCCAGCGTCCGTCACCCACTCTTCATCTG GTTGAAGATGAGGATATCCAAGCTGGTCCCATGGATGATCCTGGGGTCTC TGCTATATGTATCTATGATTTGTGTTTTCCATAGCAAATATGCAGGGTTT ATGGTCCCATACTTCCTAAGGAAATTTTTCTCCCAAAATGCCACAATTCA AAAAGAAGATACACTGGCTATACAGATTTTCTCTTTTGTTGCTGAGTTCT CAGTGCCATTGCTTATCTTCCTTTTTGCTGTTTTGCTCTTGATTTTCTCT CTGGGGAGGCACACCCGGCAAATGAGAAACACAGTGGCCGGCAGCAGGGT TCCTGGCAGGGGTGCACCCATCAGCGCGTTGCTGTCTATCCTGTCCTTCC TGATCCTCTACTTCTCCCACTGCATGATAAAAGTTTTTCTCTCTTCTCTA AAGTTTCACATCAGAAGGTTCATCTTTCTGTTCTTCATCCTTGTGATTGG TATATACCCTTCTGGACACTCTCTCATCTTAATTTTAGGAAATCCTAAAT TGAAACAAAATGCAAAAAAGTTCCTCCTCCACAGTAAGTGCTGTCAGTGA
>hGR02 aa MALSFSAILHIIMMSAEFFTGITVN GFLIIVNCNELIKHRKLMPIQILLM CIGMSRFGLQMVLMVQSFFSVFFPL LYVKIIYGAAMMFLWMFFSSISLWF ATCLSVFYCLKISGFTQSCFLWLKF RIPKLIPWLFWEAFWPL*ALHLCVE VDYAKNVEEDALRNTTLKKSKTKIK KISEVLLVNLALIFPLAIFVMCTSM LLISLYKHTHRMQHGSHGFRNANTE AHINALKTVITFFCFFISYFAAFMT NMTFSLPYRSHQFFMLKDIMAAYPS GHSVIIILSNSKFQQSFRRILCLKK KL	>hGR02 nt ATGGCCTTGTCTTTTTCAGCTATTCTTCATATTATCATGATGTCAGCAGA ATTCTTCACAGGGATCACAGTAAATGGATTTCTTATCATTGTTAACTGTA ATGAATTGATCAAACATAGAAAGCTAATGCCAATTCAAATCCTCTTAATG TGCATAGGGATGTCTAGATTTGGTCTGCAGATGGTGTTAATGGTACAAAG TTTTTTCTCTGTGTTCTTTCCACTCCTTTACGTCAAAATAATTTATGGTG CAGCAATGATGTTCCTTTGGATGTTTTTTAGCTCTATCAGCCTATGGTTT GCCACTTGCCTTTCTGTATTTTACTGCCTCAAGATTTCAGGCTTCACTCA GTCCTGTTTTCTTTGGTTGAAATTCAGGATCCCAAAGTTAATACCTTGGC TGCTTCTGGGAAGCGTTCTGGCCTCTGTGAGCATTGCATCTGTGTGTCGA GGTAGATTACGCTAAAAATGTGGAAGAGGATGCCCTCAGAAACACCACAC TAAAAAAGAGTAAAACAAAGATAAAGAAAATTAGTGAAGTGCTTCTTGTC AACTTGGCATTAATATTTCCTCTAGCCATATTTGTGATGTGCACTTCTAT GTTACTCATCTCTCTTTACAAGCACACTCATCGGATGCAACATGGATCTC ATGGCTTTAGAAATGCCAACACAGAAGCCCATATAAATGCATTAAAAACA GTGATAACATTCTTTTGCTTCTTTATTTCTTATTTTGCTGCCTTCATGAC AAATATGACATTTAGTTTACCTTACAGAAGTCACCAGTTCTTTATGCTGA AGGACATAATGGCAGCATATCCCTCTGGCCACTCGGTTATAATAATCTTG AGTAATTCTAAGTTCCAACAATCATTTAGAAGAATTCTCTGCCTCAAAAA GAAACTATGA
>hGR03 aa MMGLTEGVFLILSGTQFTLGILVNC FIELVNGSSWFKTKRMSLSDFIITT LALLRIILLCIILTDSFLIEFSPNT HDSGIIMQIIDVSWTFTNHLSIWLA TCLGVLYCLKIASFSHPTFLWLKWR VSRVMVWMLLGALLLSCGSTASLIN EFKLYSVFRGIEATRNVTEHFRKKR SEYYLIHVLGTLWYLPPLIVSLASY SLLIFSLGRHTRQMLQNGTSSRDPT TEAHKRAIRIILSFFFLFLLYFLAF LIASFGNFLPKTKMAKMIGEVMTMF YPAGHSFILILGNSKLKQTFVVMLR CESGHLKPGSKGPIFS	>hGR03 nt ATGATGGGACTCACCGAGGGGGTGTTCCTGATTCTGTCTGGCACTCAGTT CACACTGGGAATTCTGGTCAATTGTTTCATTGAGTTGGTCAATGGTAGCA GCTGGTTCAAGACCAAGAGAATGTCTTTGTCTGACTTCATCATCACCACC CTGGCACTCTTGAGGATCATTCTGCTGTGTATTATCTTGACTGATAGTTT TTTAATAGAATTCTCTCCCAACACACATGATTCAGGGATAATAATGCAAA TTATTGATGTTTCCTGGACATTTACAAACCATCTGAGCATTTGGCTTGCC ACCTGTCTTGGTGTCCTCTACTGCCTGAAAATCGCCAGTTTCTCTCACCC CACATTCCTCTGGCTCAAGTGGAGAGTTTCTAGGGTGATGGTATGGATGC TGTTGGGTGCACTGCTCTTATCCTGTGGTAGTACCGCATCTCTGATCAAT GAGTTTAAGCTCTATTCTGTCTTTAGGGGAATTGAGGCCACCAGGAATGT GACTGAACACTTCAGAAAGAAGAGGAGTGAGTATTATCTGATCCATGTTC TTGGGACTCTGTGGTACCTGCCTCCCTTAATTGTGTCCCTGGCCTCCTAC TCTTTGCTCATCTTCTCCCTGGGGAGGCACACACGGCAGATGCTGCAAAA TGGGACAAGCTCCAGAGATCCAACCACTGAGGCCCACAAGAGGGCCATCA GAATCATCCTTTCCTTCTTCTTTCTCTTCTTACTTTACTTTCTTGCTTTC TTAATTGCATCATTTGGTAATTTCCTACCAAAAACCAAGATGGCTAAGAT

	GATTGGCGAAGTAATGACAATGTTTTATG[redacted]GCTGGCCACTCATTTATTC TCATTCTGGGGAACAGTAAGCTGAAGCAGACATTTGTAGTGATGCTCCGG TGTGAGTCTGGTCATCTGAAGCCTGGATCCAAGGGACCCATTTTCTCTTA G
>hGR04 aa MLRLFYFSAIIASVILNFVGIIMNL FITVVNCKTWVKSHRISSSDRILFS LGITRFLMLGLFLVNTIYFVSSNTE RSVYLSAFFVLCFMFLDSSSVWFVT LLNILYCVKITNFQHSVFLLLKRNI SPKIPRLLLACVLISAFTTCLYITL SQASPFPELVTTRNNTSFNISEGIL SLVVSLVLSSSLQFIINVTSASLLI HSLRRHIQKMQKNATGFWNPQTEAH VGAMKLMVYFLILYIPYSVATLVQY LPFYAGMDMGTKSICLIFATLYSPG HSVLIIITHPKLKTTAKKILCFKK	>hGR04 nt ATGCTTCGGTTATTCTATTTCTCTGCTATTATTGCCTCAGTTATTTTAAA TTTTGTAGGAATCATTATGAATCTGTTTATTACAGTGGTCAATTGCAAAA CTTGGGTCAAAAGCCATAGAATCTCCTCTTCTGATAGGATTCTGTTCAGC CTGGGCATCACCAGGTTTCTTATGCTGGGACTATTTCTGGTGAACACCAT CTACTTCGTCTCTTCAAATACGGAAAGGTCAGTCTACCTGTCTGCTTTTT TTGTGTTGTGTTTCATGTTTTTGGACTCGAGCAGTGTCTGGTTTGTGACC TTGCTCAATATCTTGTACTGTGTGAAGATTACTAACTTCCAACACTCAGT GTTTCTCCTGCTGAAGCGGAATATCTCCCCAAAGATCCCCAGGCTGCTGC TGGCCTGTGTGCTGATTTCTGCTTTCACCACTTGCCTGTACATCACGCTT AGCCAGGCATCACCTTTTCCTGAACTTGTGACTACGAGAAATAACACATC ATTTAATATCAGTGAGGGCATCTTGTCTTTAGTGGTTTCTTTGGTCTTGA GCTCATCTCTCCAGTTCATCATTAATGTGACTTCTGCTTCCTTGCTAATA CACTCCTTGAGGAGACATATACAGAAGATGCAGAAAAATGCCACTGGTTT CTGGAATCCCCAGACGGAAGCTCATGTAGGTGCTATGAAGCTGATGGTCT ATTTCCTCATCCTCTACATTCCATATTCAGTTGCTACCCTGGTCCAGTAT CTCCCCTTTTATGCAGGGATGGATATGGGGACCAAATCCATTTGTCTGAT TTTTGCCACCCTTTACTCTCCAGGACATTCTGTTCTCATTATTATCACAC ATCCTAAACTGAAAACAACAGCAAAGAAGATTCTTTGTTTCAAAAAATAG
>hGR05 aa MLSAGLGLLMLVAVVEFLIGLIGNG SLVVWSFREWIRKFNWSSYNLIILG LAGCRFLLQWLIILDLSLFPLFQSS RWLRYLSIFWVLVSQASLWFATFLS VFYCKKITTFDRPAYLWLKQRAYNL SLWCLLGYFIINLLLTVQIGLTFYH PPQGNSSIRYPFESWQYLYAFQLNS GSYLPLVVFLVSSGMLIVSLYTHHK KMKVHSAGRRDVRAKAHITALKSLG CFLLLHLVYIMASPFSITSKTYPPD LTSVFIWETLMAAYPSLHSLILIMG IPRVKQTCQKILWKTVCARRCWGP	>hGR05 nt ATGCTGAGCGCTGGCCTAGGACTGCTGATGCTGGTGGCAGTGGTTGAATT TCTCATCGGTTTAATTGGAAATGGAAGCCTGGTGGTCTGGAGTTTTAGAG AATGGATCAGAAAATTCAACTGGTCCTCATATAACCTCATTATCCTGGGC CTGGCTGGCTGCCGATTTCTCCTGCAGTGGCTGATCATTTTGGACTTAAG CTTGTTTCCACTTTTCCAGAGCAGCCGTTGGCTTCGCTATCTTAGTATCT TCTGGGTCCTGGTAAGCCAGGCCAGCTTATGGTTTGCCACCTTCCTCAGT GTCTTCTATTGCAAGAAGATCACGACCTTCGATCGCCCGGCCTACTTGTG GCTGAAGCAGAGGGCCTATAACCTGAGTCTCTGGTGCCTTCTGGGCTACT TTATAATCAATTTGTTACTTACAGTCCAAATTGGCTTAACATTCTATCAT CCTCCCCAAGGAAACAGCAGCATTCGGTATCCCTTTGAAAGCTGGCAGTA CCTGTATGCATTTCAGCTCAATTCAGGAAGTTATTTGCCTTTAGTGGTGT TTCTTGTTTCCTCTGGGATGCTGATTGTCTCTTTGTATACACACCACAAG AAGATGAAGGTCCATTCAGCTGGTAGGAGGGATGTCCGGGCCAAGGCTCA CATCACTGCGCTGAAGTCCTTGGGCTGCTTCCTCTTACTTCACCTGGTTT ATATCATGGCCAGCCCCTTCTCCATCACCTCCAAGACTTATCCTCCTGAT CTCACCAGTGTCTTCATCTGGGAGACACTCATGGCAGCCTATCCTTCTCT TCATTCTCTCATATTGATCATGGGGATTCCTAGGGTGAAGCAGACTTGTC AGAAGATCCTGTGGAAGACAGTGTGTGCTCGGAGATGCTGGGGCCCATGA
>hGR06 aa MLAAALGLLMPIAGAEFLIGLVGNG VPVVCSFRGWVKKM*GVPINSHDSG K*PLSPTQADHVGHKSVSTFPEQWL ALLS*CLRVLVSQANM*FATFFSGF CCMEIMTFVXXXXXXXXXXXXXXXX XXXXLLVSFKITFYFSALVGWTL*K PLTGNSNILHPILNLLFL*IAVQ*R RLIAICDVSVPLVFL*RHHRKMEDH TAVRRRLKPRXXXXXXXXXXXXXXX LYMVSALARHFSMTF*SPSDLTILA ISATLMAVYTSFPSIVMVMRNQTCQ RIL*EMICTWKS	>hGR06 nt ATGTTGGCGGCTGCCCTAGGATTGCTGATGCCCATTGCAGGGGCTGAATT TCTCATTGGCCTGGTTGGAAATGGAGTCCCTGTGGTCTGCAGTTTTAGAG GATGGGTCAAAAAAATGTAAGGAGTCCCTATAAATTCTCATGATTCTGGT AAGTAGCCACTTTCTCCTACTCAGGCCGATCATGTTGGACATAAGTCTGT TTCCACTTTCCCAGAGCAGTGGTTGGCTTTACTATCTTAATGTCTTCGAG TCCTGGTAAGCCAGGCCAACATGTAGTTTGCCACTTTCTTCAGTGGCTTC TGCTGCATGGAGATCATGACCTTTGTCCCGCTGACTTCTTGTAGCTGAAA AGACTGGGTTTTTGTTTTTTGCTAGTGTCTTTCAAGATCACTTTTTATTT CTCAGCTCTTGTTGGCTGGACCCTTTAAAAACCCTTAACAGGAAACAGCA ACATCCTGCATCCCATTTTAAATCTGTTATTTTTATAGATTGCTGTCCAG TGAAGGAGACTGATTGCTATTTGTGATGTTTCTGTTCCACTTGTCTTTTT GTAAAGACATCACAGGAAGATGGAGGACCACACAGCTGTCAGGAGGAGGC TCAAACCAAGGTGCTCATCGCTCTGAACTTCCCCCTTTACATGGTTTCTG CCTTGGCCAGACACTTTTCCATGACCTTCTAATCTCCCTCTGATCTCACC ATTCTTGCCATCTCTGCAACACTCATGGCTGTTTATACTTCATTTCCGTC TATTGTAATGGTTATGAGGAATCAGACTTGTCAGAGAATTCTGTAGGAGA TGATATGTACATGGAAATCCTAG
>hGR07 aa MADKVQTTLLFLAVGEFSVGILGNA	>hGR07 nt ATGGCAGATAAAGTGCAGACTACTTTATTGTTCTTAGCAGTTGGAGAGTT TTCAGTGGGGATCTTAGGGAATGCATTCATTGGATTGGTAAACTGCATGG

FIGLVNCMDWVKKRKIA[redacted]LILTS LAISRICLLCVILLDCFILVLYPDV YATGKEMRIIDFFWTLTNHLSIWFA TCLSIYYFFKIGNFFHPLFLWMKWR IDRVISWILLGCVVLSVFISLPATE NLNADFRFCVKAKRKTNLTWSCRVN KTQHASTKLFLNLATLLPFCVCLMS FFLLILSLRRHIRRMQLSATGCRDP STEAHVRALKAVISFLLLFIAYYLS FLIATSSYFMPETELAVIFGESIAL IYPSSHSFILILGNNKLRHASLKVI WKVMSILKGRKFQQHKQI	ACTGGGTCAAGAAGAGGAAAATTGCCTCCA[redacted]ATTTAATCCTCACAAGT CTGGCCATATCCAGAATTTGTCTATTGTGC[redacted]ATACTATTAGATTGTTT TATATTGGTGCTATATCCAGATGTCTATGCCACTGGTAAAGAAATGAGAA TCATTGACTTCTTCTGGACACTAACCAATCATTTAAGTATCTGGTTTGCA ACCTGCCTCAGCATTTACTATTTCTTCAAGATAGGTAATTTCTTTCACCC ACTTTTCCTCTGGATGAAGTGGAGAATTGACAGGGTGATTTCCTGGATTC TACTGGGGTGCGTGGTTCTCTCTGTGTTTATTAGCCTTCCAGCCACTGAG AATTTGAACGCTGATTTCAGGTTTTGTGTGAAGGCAAAGAGGAAAACAAA CTTAACTTGGAGTTGCAGAGTAAATAAAACTCAACATGCTTCTACCAAGT TATTTCTCAACCTGGCAACGCTGCTCCCCTTTTGTGTGTGCCTAATGTCC TTTTTCCTCTTGATCCTCTCCCTGCGGAGACATATCAGGCGAATGCAGCT CAGTGCCACAGGGTGCAGAGACCCCAGCACAGAAGCCCATGTGAGAGCCC TGAAAGCTGTCATTTCCTTCCTTCTCCTCTTTATTGCCTACTATTTGTCC TTTCTCATTGCCACCTCCAGCTACTTTATGCCAGAGACGGAATTAGCTGT GATTTTTGGTGAGTCCATAGCTCTAATCTACCCCTCAAGTCATTCATTTA TCCTAATACTGGGGAACAATAAATTAAGACATGCATCTCTAAAGGTGATT TGGAAAGTAATGTCTATTCTAAAAGGAAGAAAATTCCAACAACATAAACA **AATCTGA**
>hGR08 aa MFSPADNIFIILITGEFILGILGNG YIALVNWIDWIKKKKISTVDYILTN LVIARICLISVMVVNGIVIVLNPDV YTKNKQQIVIFTFWTFANYLNMWIT TCLNVFYFLKIASSSHPLFLWLKWK IDMVVHWILLGCFAISLLVSLIAAI VLSCDYRFHAIAKHKRNITEMFHVS KIPYFEPLTLFNLFAIVPFIVSLIS FFLLVRSLWRHTKQIKLYATGSRDP STEVHVRAIKTMTSFIFFFFLYYIS SILMTFSYLMTKYKLAVEFGEIAAI LYPLGHSLILIVLNNKLRQTFVRML TCRKIACMI	>hGR08 nt **ATGTTCAGTCCTGCAGATAACATCTTTATAATCCTAATAACTGGAGAATT** CATACTAGGAATATTGGGGAATGGATACATTGCACTAGTCAACTGGATTG ACTGGATTAAGAAGAAAAAGATTTCCACAGTTGACTACATCCTTACCAAT TTAGTTATCGCCAGAATTTGTTTGATCAGTGTAATGGTTGTAAATGGCAT TGTAATAGTACTGAACCCAGATGTTTATACAAAAAATAAACAACAGATAG TCATTTTTACCTTCTGGACATTTGCCAACTACTTAAATATGTGGATTACC ACCTGCCTTAATGTCTTCTATTTTCTGAAGATAGCCAGTTCCTCTCATCC ACTTTTTCTCTGGCTGAAGTGGAAAATTGATATGGTGGTGCACTGGATCC TGCTGGGATGCTTTGCCATTTCCTTGTTGGTCAGCCTTATAGCAGCAATA GTACTGAGTTGTGATTATAGGTTTCATGCAATTGCCAAACATAAAAGAAA CATTACTGAAATGTTCCATGTGAGTAAAATACCATACTTTGAACCCTTGA CTCTCTTTAACCTGTTTGCAATTGTCCCATTTATTGTGTCACTGATATCA TTTTTCCTTTTAGTAAGATCTTTATGGAGACATACCAAGCAAATAAAACT CTATGCTACCGGCAGTAGAGACCCCAGCACAGAAGTTCATGTGAGAGCCA TTAAAACTATGACTTCATTTATCTTCTTTTTTTTCCTATACTATATTTCT TCTATTTTGATGACCTTTAGCTATCTTATGACAAAATACAAGTTAGCTGT GGAGTTTGGAGAGATTGCAGCAATTCTCTACCCCTTGGGTCACTCACTTA TTTTAATTGTTTTAAATAATAAACTGAGGCAGACATTTGTCAGAATGCTG **ACATGTAGAAAAATTGCCTGCATGATATGA**
>hGR09 aa MPSAIEAIYIILIAGELTIGIWGNG FIVLVNCIDWLKRRDISLIDIILIS LAISRICLLCVISLDGFFMLLFPGT YGNSVLVSIVNVVWTFANNSSLWFT SCLSIFYLLKIANISHPFFFWLKLK INKVMLAILLGSFLISLIISVPKND DMWYHLFKVSHEENITWKFKVSKIP GTFKQLTLNLGVMVPFILCLISFFL LLFSLVRHTKQIRLHATGFRDPSTE AHMRAIKAVIIFLLLLIVYYPVFLV MTSSALIPQGKLVLMIGDIVTVIFP SSHSFILIMGNSKLREAFLKMLRFV KCFLRRRKPFVP	>hGR09 nt **ATGCCAAGTGCAATAGAGGC**AATATATATTATTTTAATTGCTGGTGAATT GACCATAGGGATTTGGGGAAATGGATTCATTGTACTAGTTAACTGCATTG ACTGGCTCAAAAGAAGAGATATTTCCTTGATTGACATCATCCTGATCAGC TTGGCCATCTCCAGAATCTGTCTGCTGTGTGTAATATCATTAGATGGCTT CTTTATGCTGCTCTTTCCAGGTACATATGGCAATAGCGTGCTAGTAAGCA TTGTGAATGTTGTCTGGACATTTGCCAATAATTCAAGTCTCTGGTTTACT TCTTGCCTCAGTATCTTCTATTTACTCAAGATAGCCAATATATCGCACCC ATTTTTCTTCTGGCTGAAGCTAAAGATCAACAAGGTCATGCTTGCGATTC TTCTGGGGTCCTTTCTTATCTCTTTAATTATTAGTGTTCCAAAGAATGAT GATATGTGGTATCACCTTTTCAAAGTCAGTCATGAAGAAAACATTACTTG GAAATTCAAAGTGAGTAAAATTCCAGGTACTTTCAAACAGTTAACCCTGA ACCTGGGGGTGATGGTTCCCTTTATCCTTTGCCTGATCTCATTTTTCTTG TTACTTTTCTCCCTAGTTAGACACACCAAGCAGATTCGACTGCATGCTAC AGGGTTCAGAGACCCCAGTACAGAGGCCCACATGAGGGCCATAAAGGCAG TGATCATCTTTCTGCTCCTCCTCATCGTGTACTACCCAGTCTTTCTTGTT ATGACCTCTAGCGCTCTGATTCCTCAGGGAAAATTAGTGTTGATGATTGG TGACATAGTAACTGTCATTTTCCCATCAAGCCATTCATTCATTCTAATTA TGGGAAATAGCAAGTTGAGGGAAGCTTTTCTGAAGATGTTAAGATTTGTG AAGTGTTTCCTTAGAAGAAGAAAGCCTTTTGTTCCATAG
>hGR10 aa MLRVVEGIFIFVVVSESVFGVLGNG FIGLVNCIDCAKNKLSTIGFILTGL AISRIFLIWIIITDGFIQIFSPNIY	>hGR10 nt **ATGCTACGTGTAGTGGAAGG**CATCTTCATTTTTGTTGTAGTTAGTGAGTC AGTGTTTGGGGTTTTGGGGAATGGATTTATTGGACTTGTAAACTGCATTG ACTGTGCCAAGAATAAGTTATCTACGATTGGCTTTATTCTCACCGGCTTA GCTATTTCAAGAATTTTTCTGATATGGATAATAATTACAGATGGATTTAT

ASGNLIEYISYFWVI[redacted]SSMWFAT SLSIFYFLKIANFSNYIFLWLKSRT NMVLPFMIVFLLISSLLNFAYIAKI LNDYKTKNDTVWDLNMYKSEYFIKQ ILLNLGVIFFFTLSLITCIFLIISL WRHNRQMQSNVTGLRDSNTEAHVKA MKVLISFIILFILYFIGMAIEISCF TVRENKLLLMFGMTTTAIYPWGHSF ILILGNSKLKQASLRVLQQLKCCEK RKNLRVT	ACAGATATTCTCTCCAAATATATATGCC[redacted]GGTAACCTAATTGAATATA TTAGTTACTTTTGGGTAATTGGTAATCA[redacted]CAAGTATGTGGTTTGCCACC AGCCTCAGCATCTTCTATTTCCTGAAGATAGCAAATTTTTCCAACTACAT ATTTCTCTGGTTGAAGAGCAGAACAAATATGGTTCTTCCCTTCATGATAG TATTCTTACTTATTTCATCGTTACTTAATTTTGCATACATTGCGAAGATT CTTAATGATTATAAAACGAAGAATGACACAGTCTGGGATCTCAACATGTA TAAAAGTGAATACTTTATTAAACAGATTTTGCTAAATCTGGGAGTCATTT TCTTCTTTACACTATCCCTAATTACATGTATTTTTTTAATCATTTCCCTT TGGAGACACAACAGGCAGATGCAATCGAATGTGACAGGATTGAGAGACTC CAACACAGAAGCTCATGTGAAGGCAATGAAAGTTTTGATATCTTTCATCA TCCTCTTTATCTTGTATTTTATAGGCATGGCCATAGAAATATCATGTTTT ACTGTGCGAGAAAACAAACTGCTGCTTATGTTTGGAATGACAACCACAGC CATCTATCCCTGGGGTCACTCATTTATCTTAATTCTAGGAAACAGCAAGC TAAAGCAAGCCTCTTTGAGGGTACTGCAGCAATTGAAGTGCTGTGAGAAA AGGAAAAATCTCAGAGTCACATAG
>hGR11 aa MANMLKNMLTMISAIDFIMGIQRSR VMVLVHCIDWIRRWKLSLIDFILTC WAISRIFXXXXXXXXXXXXXXXXXX XXXXXXXXXXXXXXNHLCT*FATCL AVFYFLKIVNFSYLFYFWLKWRINK VAFILPLVSAFSVYQLSFDVHF*CL LVSCPKKYERHMTGLLNVSNNKNVN NIIIFFIGSLSSFSISSIFFLLLLL SS*RHMKHIRFNFRDCRTPVYGPIS EPRKRFSFFVLLLYKNLPFS	
>hGR12 aa MSSIWETLFIRILVV*FIMGTVGN* FIVLVNIID*IRN*KVSLIDFILNC LAISRICFL*ITILATSFNIGYEKM PDSKNLAVSFDILWTGSSYFCLSCT TCLSVFYFLKVANFSNPIFLWMKWK IHKVLLFIVLEATISFCTTSILKEI IINSLI*ERVTIKGNLTFNYMDTMH DFTSLFLLQMMFILPFVETLASILL LILSLWSHTRQMKLHGIYSRDPSTE AH*VKPIKAIISFLLLFIVHYFISII* *LTLACPLLD*FVAARTFSSVLVFFHP SGHSFLLILRDSKLKQASLCVLKKM KYAKKDIISHFYKHA	>hGR12 nt ATGTCAAGCATTTGGGAGACACTGTTTATAAGAATTCTTGTAGTGTAATT CATAATGGGGACTGTGGGAAATTGATTCATTGTATTGGTTAATATCATTG ACTGAATCAGGAACTGAAAGGTCTCCCTGATTGATTTTATTCTCAACTGC TTGGCCATCTCCAGGATATGTTTCCTGTAGATAACAATTTTAGCTACCTC TTTCAATATAGGCTATGAGAAAATGCCTGATTCTAAGAATCTTGCAGTAA GTTTTGACATTCTCTGGACAGGATCCAGCTATTTCTGCCTGTCCTGTACC ACTTGCCTCAGTGTCTTCTATTTCCTCAAGGTAGCCAACTTCTCCAATCC CATTTTCCTCTGGATGAAATGGAAAATTCACAAGGTGCTTCTCTTTATTG TACTAGAGGCAACGATCTCTTTCTGCACAACTTCCATTCTGAAGGAAATA ATAATTAATAGTTTAATCTAAGAACGGGTAACAATAAAAGGCAACTTGAC ATTTAATTATATGGATACCATGCATGATTTCACTTCTCTGTTTCTCCTTC AGATGATGTTCATCCTTCCTTTTGTGGAAACACTGGCTTCCATTCTTCTC TTAATCCTCTCCTTATGGAGCCACACCAGGCAGATGAAGCTACATGGTAT TTATTCCAGGGATCCCAGCACAGAAGCCCAT*GTAAAACCTATAAAAGCTA* *TAATTTCATTTCTACTCCTCTTTATTGTGCATTATTTCATCAGTATCATA* *CTAACATTGGCCTGTCCTCTTCTAG*ACTTCGTTGCGGCAAGGACTTTTAG TAGTGTGCTGGTATTTTTCCATCCATCTGGCCATTCATTTCTTCTAATTT TACGGGACAGCAAACTGAAGCAAGCTTCTCTCTGTGTCCTGAAGAAGATG AAGTATGCCAAAAAGGACATAATCTCTCATTTTTATAAACATGCCTGA
>hGR13 aa MESALPSIFTLVIIAEFIIGNLSNG FIVLINCIDWVSKRELSSVDKLLII LAISRIGLIWEILVSWFLALHYLAI FVSGTGLRIMIFSWIVSNHFNLWLA TIFSIFYLLKIASFSSPAFLYLKWR VNKVILMILLGTLVFLFLNLIQINM HIKDWLDRYERNTTWNFSMSDFETF SVSVKFTMTMFSLTPFTVAFISFLL LIFSLQKHLQKMQLNYKGHRDPRTK VHTNALKIVISFLLFYASFFLCVLI SWISELYQNTVIYMLCETIGVFSPS	>hGR13 nt ATGGAAAGTGCCCTGCCGAGTATCTTCACTCTTGTAATAATTGCAGAATT CATAATTGGGAATTTGAGCAATGGATTTATAGTACTGATCAACTGCATTG ACTGGGTCAGTAAAAGAGAGCTGTCCTCAGTCGATAAACTCCTCATTATC TTGGCAATCTCCAGAATTGGGCTGATCTGGGAAATATTAGTAAGTTGGTT TTTAGCTCTGCATTATCTAGCCATATTTGTGTCTGGAACAGGATTAAGAA TTATGATTTTTAGCTGGATAGTTTCTAATCACTTCAATCTCTGGCTTGCT ACAATCTTCAGCATCTTTTATTTGCTCAAAATAGCGAGTTTCTCTAGCCC TGCTTTTCTCTATTTGAAGTGGAGAGTAAACAAAGTGATTCTGATGATAC TGCTAGGAACCTTGGTCTTCTTATTTTTAAATCTGATACAAATAAACATG CATATAAAAGACTGGCTGGACCGATATGAAAGAAACACAACTTGGAATTT CAGTATGAGTGACTTTGAAACATTTTCAGTGTCGGTCAAATTCACTATGA CTATGTTCAGTCTAACACCATTTACTGTGGCCTTCATCTCTTTTCTCCTG

SHSFLLILGNAKLRQ[redacted]LVAAKVW AKR	TTAATTTTCTCCCTGCAGAAACATCTCC[redacted]AAATGCAACTCAATTACAA AGGACACAGAGACCCCAGGACCAAGGTC[redacted]ACAAATGCCTTGAAAATTG TGATCTCATTCCTTTTATTCTATGCTAGTTTCTTTCTATGTGTTCTCATA TCATGGATTTCTGAGCTGTATCAGAACACAGTGATCTACATGCTTTGTGA GACGATTGGAGTCTTCTCTCCTTCAAGCCACTCCTTTCTTCTGATTCTAG GAAACGCTAAGTTAAGACAGGCCTTTCTTTTGGTGGCAGCTAAGGTATGG GCTAAACGATGA
>hGR14 aa MGGVIKSIFTFVLIVEFIIGNLGNS FIALVNCIDWVKGRKISSVDRILTA LAISRISLVWLIFGSWCVSVFFPAL FATEKMFRMLTNIWTVINHFSVWLA TGLGTFYFLKIANFSNSIFLYLKWR VKKVVLVLLLVTSVFLFLNIALINI HINASINGYRRNKTCSSDSSNFTRF SSLIVLTSTVFIFIPFTLSLAMFLL LIFSMWKHRKKMQHTVKISGDASTK AHRGVKSVITFFLLYAIFSLSFFIS VWTSERLEENLIILSQVMGMAYPSC HSCVLILGNKKLRQASLSVLLWLRY MFKDGEPSGHKEFRESS	>hGR14 nt ATGGGTGGTGTCATAAAGAGCATATTTACATTCGTTTTAATTGTGGAATT TATAATTGGAAATTTAGGAAATAGTTTCATAGCACTGGTGAACTGTATTG ACTGGGTCAAGGGAAGAAAGATCTCTTCGGTTGATCGGATCCTCACTGCT TTGGCAATCTCTCGAATTAGCCTGGTTTGGTTAATATTCGGAAGCTGGTG TGTGTCTGTGTTTTTCCCAGCTTTATTTGCCACTGAAAAAATGTTCAGAA TGCTTACTAATATCTGGACAGTGATCAATCATTTTAGTGTCTGGTTAGCT ACAGGCCTCGGTACTTTTTATTTTCTCAAGATAGCCAATTTTTCTAACTC TATTTTTCTCTACCTAAAGTGGAGGGTTAAAAAGGTGGTTTTGGTGCTGC TTCTTGTGACTTCGGTCTTCTTGTTTTTAAATATTGCACTGATAAACATC CATATAAATGCCAGTATCAATGGATACAGAAGAAACAAGACTTGCAGTTC TGATTCAAGTAACTTTACACGATTTTCCAGTCTTATTGTATTAACCAGCA CTGTGTTCATTTTCATACCCTTTACTTTGTCCCTGGCAATGTTTCTTCTC CTCATCTTCTCCATGTGGAAACATCGCAAGAAGATGCAGCACACTGTCAA AATATCCGGAGACGCCAGCACCAAAGCCCACAGAGGAGTTAAAAGTGTGA TCACTTTCTTCCTACTCTATGCCATTTTCTCTCTGTCTTTTTTCATATCA GTTTGGACCTCTGAAAGGTTGGAGGAAAATCTAATTATTCTTTCCCAGGT GATGGGAATGGCTTATCCTTCATGTCACTCATGTGTTCTGATTCTTGGAA ACAAGAAGCTGAGACAGGCCTCTCTGTCAGTGCTACTGTGGCTGAGGTAC ATGTTCAAAGATGGGGAGCCCTCAGGTCACAAAGAATTTAGAGAATCATC TTGA
>hGR15 aa MITFLPIIFSILVVVTFVLGNFANG FIVLVNSIEWVKRQKISFADQILTA LAVSRVGLLWVILLHWYATVLNPGS YSLGVRITTINAWAVTNHFSIWVAT SLSIFYFLKIANFSNFIFLHLKRRI KSVIPVILLGSLLFLVCHLVVVNMD ESMWTKEYEGNVSWEIKLSDPTHLS DMTVTTLANLIPFTLSLLSFLLLIC SLCKHLKKMQFHGKGSPDSNTKVHI KALQTVTSFLLLFAVYFLSLITSIW NFRRRL*NEPVLMLSQTTAIIYPSF HSFILIWGSKKLKQTFLLILCQIKC	>hGR15 nt ATGATAACTTTTCTACCCATCATTTTTTCCATTCTAGTAGTGGTTACATT TGTTCTTGGGAATTTTGCTAATGGCTTCATAGTGTTGGTAAATTCCATTG AGTGGGTCAAGAGACAAAAGATCTCCTTTGCTGACCAAATTCTCACTGCT CTGGCAGTCTCCAGAGTTGGTTTGCTCTGGGTAATATTATTACATTGGTA TGCAACTGTTTTGAATCCAGGTTCATATAGTTTAGGAGTAAGAATTACTA CTATTAATGCCTGGGCTGTAACCAACCATTTCAGCATCTGGGTTGCTACT AGCCTCAGCATATTTTATTTCCTCAAGATTGCCAATTTCTCCAACTTTAT TTTTCTTCACTTAAAAAGGAGAATTAAGAGTGTCATTCCAGTGATACTAT TGGGGTCTTTGTTATTTTTGGTTTGTCATCTTGTTGTGGTAAACATGGAT GAGAGTATGTGGACAAAAGAATATGAAGGAAACGTGAGTTGGGAGATCAA ATTGAGTGATCCGACGCACCTTTCAGATATGACTGTAACCACGCTTGCAA ACTTAATACCCTTTACTCTGTCCCTGTTATCTTTTCTGCTCTTAATCTGT TCTTTGTGTAAACATCTCAAGAAGATGCAGTTCCATGGCAAAGGATCTCC AGATTCCAACACCAAGGTCCACATAAAAGCTTTGCAAACGGTGACCTCCT TCCTCTTGTTATTTGCTGTTTACTTTCTGTCCCTAATCACATCGATTTGG AATTTTAGGAGGAGGCTGTAGAACGAACCTGTCCTCATGCTCAGCCAAAC TACTGCAATTATATACCCTTCATTTCATTCATTCATCCTAATTTGGGGAA GCAAGAAGCTGAAACAGACCTTTCTTTTGATTTTGTGTCAGATTAAGTGC TGA
>hGR16 aa MIPIQLTVFFMIIYVLESLTIIVQS SLIVAVLGREWLQVRRLMPVDMILI SLGISRFCLQWASMLNNFCSYFNLN YVLCNLTITWEFFNILTFWLNSLLT VFYCIKVSSFTHHIFLWLRWRILRL FPWILLGSLMITCVTIIPSAIGNYI QIQLLTMEHLPRNSTVTDKLENFHQ YQFQAHTVALVIPFILFLASTIFLM ASLTKQIQHHSTGHCNPSMKARFTA LRSLAVLFIVFTSYFLTILITIIGT LFDKRCWLWVWEAFVYAFILMHSTS LMLSSPTLKRILKGKC	>hGR16 nt ATGATACCCATCCAACTCACTGTCTTCTTCATGATCATCTATGTGCTTGA GTCCTTGACAATTATTGTGCAGAGCAGCCTAATTGTTGCAGTGCTGGGCA GAGAATGGCTGCAAGTCAGAAGGCTGATGCCTGTGGACATGATTCTCATC AGCCTGGGCATCTCTCGCTTCTGTCTACAGTGGGCATCAATGCTGAACAA TTTTTGCTCCTATTTTAATTTGAATTATGTACTTTGCAACTTAACAATCA CCTGGGAATTTTTTAATATCCTTACATTCTGGTTAAACAGCTTGCTTACC GTGTTCTACTGCATCAAGGTCTCTTCTTTCACCCATCACATCTTTCTCTG GCTGAGGTGGAGAATTTTGAGGTTGTTTCCCTGGATATTACTGGGTTCTC TGATGATTACTTGTGTAACAATCATCCCTTCAGCTATTGGGAATTACATT CAAATTCAGTTACTCACCATGGAGCATCTACCAAGAAACAGCACTGTAAC TGACAAACTTGAAAATTTTCATCAGTATCAGTTCCAGGCTCATACAGTTG CATTGGTTATTCCTTTCATCCTGTTCCTGGCCTCCACCATCTTTCTCATG GCATCACTGACCAAGCAGATACAACATCATAGCACTGGTCACTGCAATCC AAGCATGAAAGCGCGCTTCACTGCCCTGAGGTCCCTTGCCGTCTTATTTA

	TTGTGTTTACCTCTTACTTTCTAACCAT[redacted]CATCACCATTATAGGTACT CTATTTGATAAGAGATGTTGGTTATGGGT[redacted]GGGAAGCTTTTGTCTATGC TTTCATCTTAATGCATTCCACTTCACTGATGCTGAGCAGCCCTACGTTGA AAAGGATTCTAAAGGGAAAGTGCTAG
>hGR17 aa MCSAXLLIILSILVVFAFVLGNVAN GFIALINVNDWVKTQKISSTDQIVT ALAFSRIGLLXTLIILLHWYATVFN SALYSLEVRIVPSNVSAIINHFSIW LATSLSIFYLFKIANFSNFIFLHLK KRIKSVLLVILLGSLVFLICNLAVV TMDDSVWTKEFEGNVTWKIELRNAI HLSNMTITNHASKLHTVHSDSNIFS AVSLFSXTMLANFTLFILTLISFLL LVCSPCKHLKMMQLHGKGSQDLSTK VHIKPLQTVISFRMLFAIYFLCIIT STWNPRTQQSNLVFLLYQTLAIMYP SFHSFILIMRSRKLKQTSLSVLCQV TCWVK	>hGR17 nt
>hGR18 aa MFVGINIFFLVVATRGLVLGMLGNG LIGLVNCIEWAKSWKVSSADFILTS LAIVRIIRLYLILFDSFIMVLSPHL YTIRKLVKLFTILWALINQLSI*FA TCLSIFYLLKIANFSHSLFLWLKWR MNGMIVMLLILSLFLLIFDSLVLEI FIDISLNIIDKSNLTLYLDESKTLY DKLSILKTLLSLTYVIPFLLTLTSL LLLFISLVRHTKNLQLNSLGSRDSS TEAHKRAMKMVIAFLLLFIINFIST LIGDWIFLEVENYQVMMFIMMILLA FPSGHSFIIILGNNKLRQSSLRLLW HLKFSLKKAKPLTS	>hGR18 nt ATGTTCGTTGGAATTAATATTTTCTTTCTGGTGGTGGCAACAAGAGGACT TGTCTTAGGAATGCTGGGAAACGGGCTCATTGGACTGGTAAACTGCATTG AGTGGGCCAAGAGTTGGAAGGTCTCATCAGCTGATTTCATCCTCACCAGC TTGGCTATAGTCAGAATCATTCGACTGTATTTAATACTATTTGATTCATT TATAATGGTATTGTCCCCTCATCTATATACCATCCGTAAACTAGTAAAAC TGTTTACTATTCTTTGGGCATTAATTAATCAGTTAAGTATCTAGTTTGCC ACCTGCCTAAGCATTTTCTACTTGCTTAAGATAGCCAATTTCTCCCACTC CCTTTTCCTCTGGCTGAAGTGGAGAATGAACGGAATGATTGTTATGCTTC TTATATTGTCTTTGTTCTTACTGATTTTTGACAGTTTAGTGCTAGAAATA TTTATTGATATCTCACTCAATATAATAGATAAAAGTAATCTGACTTTATA TTTAGATGAAAGTAAAACTCTCTATGATAAACTCTCTATTTTAAAAACTC TTCTCAGCTTGACATACGTTATTCCCTTTCTTCTGACTCTGACCTCTTTG CTCCTTTTATTTATATCCTTAGTGAGACACACCAAGAATTTGCAGCTCAA CTCTCTGGGCTCAAGGGACTCCAGCACAGAGGCCCATAAAAGGGCCATGA AAATGGTGATAGCCTTCCTCCTCCTTTTTATTATTAACTTTATTTCCACT TTAATAGGAGATTGGATCTTCCTTGAGGTAGAGAATTATCAGGTCATGAT GTTTATTATGATGATTTTACTTGCCTTTCCCTCAGGCCACTCATTTATTA TAATTTTGGGAAACAACAAGCTAAGACAGAGCTCCTTGAGACTACTGTGG CATCTTAAATTCTCTCTGAAAAAAGCAAAACCTTTAACTTCATAG
>hGR19 aa VTTLANLIPFTLSLICFLLLICSLC KHLKKMRLHSKGSQDPSTKVHIKAL QTVTSFLMLFAIYFLCIITSTWNLR TQQSKLVLLLCQTVAIMYPSFHSFI LIMGSRKLKQTFLSVLWQMTC	>hGR19 nt CTGTAACTACTCTAGCAAACCTCATACCCTTTACTCTGAGCCTAATATGTT TTCTGCTGTTAATCTGTTCTCTTTGTAAACATCTCAAGAAGATGCGGCTCC ATAGCAAAGGATCTCAAGATCCCAGCACCAAGGTCCATATAAAAGCTTTGC AAACTGTGACCTCCTTCCTCATGTTATTTGCCATTTACTTTCTGTGTATAA TCACATCAACTTGGAATCTTAGGACACAGCAGAGCAAACTTGTACTCCTGC TTTGCCAAACTGTTGCAATCATGTATCCTTCATTCCACTCATTCATCCTGA TTATGGGAAGTAGGAAGCTAAAACAGACCTTTCTTTCAGTTTTGTGGCAGA TGACATGCTGAGTGAAAGAAGAGAAACCCTCAACTCCATAGATTCACAAGG GGAGCATCGTGGGTCTTCTAGCAGAAAACAAACTGATGGTGTCTGGAACAT TTTATAT
>hGR20 aa HLXRKAKSVVLVIVLGSLFFLVCQL VMKNTYINVWTEECEGNVTWKIKLR NAMHLSNLTVAMLANLIPFTLTVIS FLLLIYSLCKHLKKMQLHGKGSQDP STKIHIKALQTVTSFLVLLAIYFLC LIIS	>hGR20 nt TTCATCACTTANAAAGGAAGGCTAAGAGTGTAGTTCTGGTGATAGTGTTG GGGTCTTTGTTCTTTTTGGTTTGTCAACTTGTGATGAAAAACACGTATAT AAATGTGTGGACAGAAGAATGTGAAGGAAACGTAACTTGGAAGATCAAAC TGAGGAATGCAATGCACCTTTCCAACTTGACTGTAGCCATGCTAGCAAAC TTGATACCATTCACTCTGACCGTGATATCTTTTCTGCTGTTAATCTACTC TCTGTGTAAACATCTGAAGAAGATGCAGCTCCATGGCAAAGGATCTCAAG ATCCCAGCACCAAGATCCACATAAAAGCTCTGCAAACTGTGACCTCCTTC CTCGTATTACTTGCCATTTACTTTCTGTGTCTAATCATATCCTTTTG
>hGR21 aa MPPGIGNTFLIVMMGEFII*MLGNG	

FIVLVNCIDW*GVK*S TTASSPA WLSPQSVNFG*YYLIHL*QHYGHIY MPSIN**NLFIFFGH*PIT*LPGLL P*CFLLL*NTYFSHPCFIWLRWRIS RTLLELPLGSLLLLFFNLALTGGLS DLWINIYTIYERNSTWSLDVSKILY CSLWILVSLIYLISFLLSLISLLLL ILSLMRHIRNLQLNTMGPRDLRMKA HKRAMKMKMKMMVSFLLFFLVHFSS LLPTGWIFLIQQK*QANFFVLLTSI IFPSSHSFVLILENCKLRQTAVGPL WHLKCHLKRVKL	
>hGR22 aa MATESDTNLLILAIAEFIISMLGNV FIGLVNCSEXIKNXKVFSADFILTC LAISHNGQLLVILFDSFLVGLASHL YTTYRLXKNCIMLWT	>hGR22 nt TATAGGGACNGTGATGCTTCGTACACTCTCCAAGAAGAAACACTCCGTGAG GTATGTGAGACTGCATNCCTTAGTAGATCTNTTGGGATATATATTCATAAT ATAGAAAAANAGGCAAAGACTTNCTTAAGTATATGAGACTCTATCCAACAG CAGAAGGTTCTGATCAAGACTGGAAGTGCAATANAAGCAATGAAGATAAGT ATCAGATATGAATGCTCTTCTGCAATGGTCTGATTGTNACATTATTAATGA TACANAGTATTAAAAACTTGGATTTTNTTGTCTCTGGAGATGGCCACCGAA TCGGACACAAATCTTCTGATTCTGGCAATAGCAGAATTCATCATCAGCATG CTGGGGAATGTGTTCATTGGACTGGTAAACTGCTCTGAANGGATCAAGAAC CANAAGGTCTTCTCAGCTGACTTCATCCTCACCTGCTTGGCTATCTCTCAC AATGGACAACTGTTGGTGATACTGTTTGATTCATTTCTAGTGGGACTTGCT TCACATCTATATACCACATATAGACTANGAAAAAACTGTATTATGCTTTGG ACATGACTAATCACTTGACACACTGCTTCGCACGTGCTAGCATATTCTATT CTTAGATAGCCACTTCNCACTCCTTGTCTCTGCTGAAGTGGGAT
>hGR23 aa VAFVLGNVANGFIALVNVIDXVNTR KISSAEQILTALVVSRIGXTLXHSI P*DATRC*SALYRXEVRIVASN	>hGR23 nt AGGGTTGAGTCGTGCTTATCTTCACTTAACCTAGTATANAANTACAGCATA TAGCAAGGAGAGAATGTATATGAAGAGGAGTGAATTTGAGTCTGTTTGAGA ATAATGACCTTTTCTATTTCTATAAAGACAGTTTTGAATTCATCTATTAGC ATATGCTGGTGCTTGCCTGTTGACACTAGTCACTGAATTTAAAGGCAGAAA ATGTTATTGCACATTTAGTAATCAAGTGTTCATCGAAGTTAACATCTGGAT GTTAAAGGACTCAGAACAAGTGTTACTAAGCCTGCATTTTTTTATCTGTTC AAACATGATGTGTTNTCTGCTCATCATTTCATCAATTCTGGTAGAGTTGCA TTTGTTCTTGGAAATGTNGCCAATGGCTTCATAGCTCTAGTAAATGTCATT GACTGNGTTAACACACGAAAGATCTCCTCAGCTGAGCAAATTCTCACTGCT CTGGTGGTCTCCAGAATTGGTNNTACTCTGNGTCATAGTATTCCTTGAGAT GCAACTAGATGTTAATCTGCTCTATATAGGNTAGAAGTAAGAATTGTTGCT TCTAATGCCTGAGCTCGTACGAACCATT
>hGR24 aa MATELDKIFLILAIAEFIISMLGNV FIGLVNCSEGIKNQKVFSADFILTC LAISTIGQLLVILFDSFLVGLASHL YTTYRLGKTVIMLWHMTNHLTTWLA TCLSIFYFFKIAHFPHSLFLWLRWR MNGMIVMLLILSLFLLIFDSLVLEI FIDISLNIIDKSNLTLYLDESKTLY DKLSILKTLLSLTSFIPFSLFLTSL LFLFLSLVRHTRNLKLSSLGSRDSS TEAHRRAMKMVMSFLFLFIVHFFSL QVANGIFFMLWNNKYIKFVMLALNA FPSCHSFILILGNSKLRQTAVRLLW HLRNYTKTPNALPL	>hGR24 nt ATGGCCACCGAATTGGACAAAATCTTTCTGATTCTGGCAATAGCAGAATTC ATCATCAGCATGCTGGGGAATGTGTTCATTGGACTGGTAAACTGCTCTGAA GGGATCAAGAACCAAAAGGTCTTCTCAGCTGACTTCATCCTCACCTGCTTG GCTATCTCCACAATTGGACAACTGTTGGTGATACTGTTTGATTCATTTCTA GTGGGACTTGCTTCACATTTATATACCACATATAGACTAGGAAAAACTGTT ATTATGCTTTGGCACATGACTAATCACTTGACAACCTGGCTTGCCACCTGC CTAAGCATTTTCTATTTCTTTAAGATAGCCCACTTCCCCCACTCCCTTTTC CTCTGGCTGAGGTGGAGGATGAACGGAATGATTGTTATGCTTCTTATATTG TCTTTGTTCTTACTGATTTTTGACAGTTTAGTGCTAGAAATATTTATTGAT ATCTCACTCAATATAATAGATAAAAGTAATCTGACTTTATATTTAGATGAA AGTAAAACTCTCTATGATAAACTCTCTATTTTAAAAACTCTTCTCAGCTTA ACCAGTTTTATCCCCTTTTCTCTGTTCCTGACCTCCTTGCTTTTTTTATTT CTGTCCTTGGTGAGACATACTAGAAATTTGAAGCTCAGTTCCTTGGGCTCT AGAGACTCCAGCACAGAGGCCCATAGGAGGGCCATGAAAATGGTGATGTCT TTCCTTTTCCTCTTCATAGTTCATTTTTTTTCCTTACAAGTGGCCAATGGG ATATTTTTTATGTTGTGGAACAACAAGTACATAAAGTTTGTCATGTTAGCC TTAAATGCCTTTCCCTCGTGCCACTCATTTATTCTCATTCTGGGAAACAGC AAGCTGCGACAGACAGCTGTGAGGCTACTGTGGCATCTTAGGAACTATACA AAAACACCAAATGCTTTACCTTTGTAG

>hGR25 aa LSPFRMLFAIYFLCIITSTWNPRTQ QSNLVFLLYQTLAIMYPSFHSFILI MRSRKLKQTSLSVLCQVTCWVK	>hGR25 nt
>hGR26 aa MPPGIGNTFLIVMMGEFII*MLGNG FIVLVNCIDVRSQMILLDNCILTSL AISTISQLWIILLDSFVTALWPHLY AFNKLIKFIHIFWALTNHLVTWLAC CLSVFYFFKIAYFSHPCFIWLRWRI SRTLLELPLGSLLLLFFNLALTGGL SDLWINIYTMYERNSTWSLDVSKIL YCSLWILVSLIYLISFLLSLISLLL LILSLMRHIRNLQLNTMGPRDLRMK AHKRAMKMKMKMMVSFLLFFLVHFS SLLPTGWIFLIQQK	>hGR26 nt
>hGR27 aa LANLIDWAENQICLMDFILSSLAIC RTLLLGCCVAIRCTYNDYPNIDAVN HNLIKIITIFDILRLVSK*LGIWFA SYLSIFYLLKVALFHHAIFLWLKWR ISRAVFTFLMIFLFFYISIISMIKI KLFLDQC*YKI*EKLLLEGRCE*SP PSC*PDAH*PGVVYSLYHFSYLMFL VCYLPKGKHCTAVVIGDWLQRPRTE AYVRAMNIMIAFFFHLLYSLGTSLS SVSYFLCKRKIVALGAYLSYPLSHS FILIMENNKVRKAL	
>hGR28 aa NICVLLIILSILVVSAFVLGNVANG FIALINVNDW	>hGR28 nt
>hGR29 aa MQAALTAFFVLLFSLLSLLGIAANG FIVLVLGKEWL	>hGR29 nt
>hGR30 aa MITFLPIIFSILVVVTFVLGNFSNG FIALVNSIEWVKTRKISSADQILTA LVVSRVGLLWVILLHWYANVFNSAL YSSEVGAVASNISAIINHFSIWLAT SLSIFYLLKIANFSNLIFLHLKKRI RSVVLVILLGPLVFLICNLAVITMD DSVWTKEYEGNVTWKIKLRNAIHLS NMTVSTLANLIPFILTLICFLLLIC SLCKHLKKMQLHGKGSQDPSTKVHI KALQTVTSFLLLCAIYFLSMIISVC NFGRLEKQPVFMFCQAIIFSYPSTH PFILILGNKKLKQIFLSVLRHVRYW VKDRSLRLHRFTRGALCVF	>hGR30 nt ATGATAACTTTTCTACCCATCATTTTTTCCATTCTGGTAGTGGTTACATTT GTTCTTGGAAATTTTTCCAATGGCTTCATAGCTCTAGTAAATTCCATTGAG TGGGTCAAGACACGAAAGATCTCCTCAGCTGACCAAATCCTCACTGCTCTG GTGGTCTCCAGAGTTGGTTTACTCTGGGTCATATTATTACATTGGTATGCA AATGTGTTTAATTCAGCTTTATATAGTTCAGAAGTAGGAGCTGTTGCTTCT AATATCTCAGCAATAATCAACCATTTCAGCATCTGGCTTGCTACTAGCCTC AGCATATTTTATTTGCTCAAGATTGCCAATTTCTCCAACCTTATTTTTCTC CACTTAAAGAAGAGAATTAGGAGTGTTGTTCTGGTGATACTGTTGGGTCCC TTGGTATTTTTGATTTGTAATCTTGCTGTGATAACCATGGATGACAGTGTG TGGACAAAAGAATATGAAGGAAATGTGACTTGGAAGATCAAATTGAGGAAT GCAATACACCTTTCAAATATGACTGTAAGCACACTAGCAAACCTCATACCC TTCATTCTGACCCTAATATGTTTTCTGCTGTTAATCTGTTCTCTGTGTAAA CATCTCAAGAAGATGCAGCTCCATGGCAAAGGATCTCAAGATCCCAGCACC AAGGTCCACATAAAAGCTTTGCAAACTGTGACCTCCTTTCTTCTGTTATGT GCCATTTACTTTCTGTCCATGATCATATCAGTTTGTAATTTTGGGAGGCTG GAAAAGCAACCTGTCTTCATGTTCTGCCAAGCTATTATATTCAGCTATCCT TCAACCCACCCATTCATCCTGATTTTGGGAAACAAGAAGCTAAAGCAGATT TTTCTTTCAGTTTTGCGGCATGTGAGGTACTGGGTGAAAGACAGAAGCCTT CGTCTCCATAGATTCACAAGAGGGGCATTGTGTGTCTTCTAG
>hGR31 aa MTTFIPIIFSSVVVVLFVIGNFANG FIALVNSIERVKRQKISFADQILTA	>hGR31 nt ATGACAACTTTTATACCCATCATTTTTTCCAGTGTGGTAGTGGTTCTATT TGTTATTGGAAATTTTGCTAATGGCTTCATAGCATTGGTAAATTCCATTG AGCGGGTCAAGAGACAAAAGATCTCTTTTGCTGACCAGATTCTCACTGCT

LAVSRVGLLWVLLLN[redacted]TVFNPAF YSVEVRTTAYNVWAVTGHFSNWLAT SLSIFYLLKIANFSNLIFLHLKRRV KSVILVMLLGPLLFLACQLFVINMK EIVRTKEFEGNMTWKIKLKSAMYFS XMTVTIGAXLVPFTLSLISFLMLIC SLCKHLKKMQLHGEGSQDLSTKVHI KALQTLISFLLLCAIFFLFLIVSVW SPRRLRNDPVVMVSKAVGNIYLAFD SFILIWRTKKLKHTFLLILCQIRC	CTGGCGGTCTCCAGAGTTGGTTTGCTCT[redacted]TATTATTATTAAATTGGTA TTCAACTGTGTTTAATCCAGCTTTTTATAGTGTAGAAGTAAGAACTACTG CTTATAATGTCTGGGCAGTAACCGGCCATTTCAGCAACTGGCTTGCTACT AGCCTCAGCATATTTTATTTGCTCAAGATTGCCAATTTCTCCAACCTTAT TTTTCTTCACTTAAAGAGGAGAGTTAAGAGTGTCATTCTGGTGATGCTGT TGGGGCCTTTACTATTTTTGGCTTGTCAACTTTTTGTGATAAACATGAAA GAGATTGTACGGACAAAAGAATTTGAAGGAAACATGACTTGGAAGATCAA ATTGAAGAGTGCAATGTACTTTTCANATATGACTGTAACCATTGGAGCAN ACTTAGTACCCTTTACTCTGTCCCTGATATCTTTTCTGATGCTAATCTGT TCTCTGTGTAAACATCTCAAGAAGATGCAGCTCCATGGAGAAGGATCGCA AGATCTCAGCACCAAGGTCCACATAAAAGCTTTGCAAACTCTGATCTCCT TCCTCTTGTTATGTGCCATTTTCTTTCTATTCCTAATCGTTTCGGTTTGG AGTCCTAGGAGGCTGCGGAATGACCCGGTTGTCATGGTTAGCAAGGCTGT TGGAAACATATATCTTGCATTCGACTCATTCATCCTAATTTGGAGAACCA AGAAGCTAAAACACACCTTTCTTTTGATTTTGTGTCAGATTAGGTGCTGA
>hGR32 aa HSFMLTMGSRKPKQTFLSAL	
>hGR33 aa MVYFLPIIFSILVVFAFVLGNFSNG FIALVNVIDWVKRQKISSADQILTA LVVSRVGLLWVILLHWYANVFNSAL YSLEVRIVASNISAVINHFSIWLAA SLSIFYLLKIANFSNLIFLHLKKRI KSVVLVILLGPLVFLICNLAVITMD ERVWTKEYEGNVTWKIKLRNAIHLS SLTVTTLANLIPFTLSLICFLLLIC SLCKHLKKMQLHSKGSQDPSTKVHI KALQTVISFLMLCAIYFLSIMISVW NLRSLENKPVFMFCKAIRFSYPSIH PFILIWGNKKLKQTFLSVFWQVRYW VKGEKPSSP	>hGR33 nt ATGGTATATTTTCTGCCCATCATTTTTTCCATTCTGGTAGTGTTTGCATT TGTTCTTGGAAATTTTTCCAATGGCTTCATAGCTCTAGTAAATGTCATTG ACTGGGTTAAGAGACAAAAGATCTCCTCAGCTGACCAAATTCTCACTGCT CTGGTGGTCTCCAGAGTTGGTTTACTCTGGGTCATATTATTACATTGGTA TGCAAATGTGTTTAATTCAGCTTTATATAGTTTAGAAGTAAGAATTGTTG CTTCTAATATCTCAGCAGTAATCAACCATTTCAGCATCTGGCTTGCTGCT AGCCTCAGCATATTTTATTTGCTCAAGATTGCCAATTTCTCCAACCTTAT TTTTCTCCACCTAAAGAAGAGAATTAAGAGTGTTGTTCTGGTGATACTGT TGGGGCCCTTGGTATTTCTGATTTGTAATCTTGCTGTGATAACCATGGAT GAGAGAGTGTGGACAAAAGAATATGAAGGAAATGTGACTTGGAAGATCAA ATTGAGGAATGCAATACACCTTTCAAGCTTGACTGTAACTACTCTAGCAA ACCTCATACCCTTTACTCTGAGCCTAATATGTTTTCTGCTGTTAATCTGT TCTCTTTGTAAACATCTCAAGAAGATGCAGCTCCATAGCAAAGGATCTCA AGATCCCAGCACCAAGGTCCACATAAAAGCTTTGCAAACTGTGATCTCCT TCCTCATGTTATGTGCCATTTACTTTCTGTCCATAATGATATCAGTTTGG AATCTTAGGAGTCTGGAAAACAAACCTGTCTTCATGTTCTGCAAAGCTAT TAGATTCAGCTATCCTTCAATCCACCCATTCATCCTGATTTGGGGAAACA AGAAGCTAAAGCAGACTTTTCTTTCAGTTTTTTGGCAAGTGAGGTACTGG GTGAAAGGAGAGAAGCCTTCATCTCCATAG
>hGR34 aa GSSRXKPPRIPHKKLCKLGPSFPHN NLPIYFLCXNHIVLEFLKMRPKKKC SLMLCQAFGIIYPSFHSFILXWGNK TLKQTFLSVXWQVTCWAKGQNQSTP	
>hGR35 aa NAIRPSKLWTVTEADKTSQPGTSANK FSAGNLISHVNMSRRMQLHGKGSQHL TRVHIKAXQTVISFLMLXAIYFLCLI STWNPRTQQSKLVFLLYQTLGFMYLL HSFILTMGSRKPKQTFLSAL	
>hGR36 aa MICFLLIILSILVVFAFVLGNFSNG FIALVNVIDWVKRQKISSADQILTA LVVSRVGLLWVILLHWYSNVLNSAL YSSEVIIFISNAWAIINHFSIWLAT SLSIFYLLKIVNFSRLIFHHLKRKA KSVVLVIVLGPLVFLVCHLVMKHTY INVWTKEYEGNVTWKIKLRNAIHLS NLTVSTLANLIPFTLTLISFLLLIY SLCKHLKKMQLHGKGSQDPSTKVHI	>hGR36 nt ATGATATGTTTTCTGCTCATCATTTTATCAATTCTGGTAGTGTTTGCATT TGTTCTTGGAAATTTTTCCAATGGCTTCATAGCTCTAGTAAATGTCATTG ACTGGGTCAAGAGACAAAAGATCTCCTCAGCTGACCAAATCCTCACTGCT CTGGTGGTCTCCAGAGTTGGTTTACTCTGGGTAATATTATTACATTGGTA TTCAAATGTGTTGAATTCAGCTTTATATAGTTCAGAAGTAATAATTTTTA TTTCTAATGCCTGGGCAATAATCAACCATTTCAGCATCTGGCTTGCTACT AGCCTCAGCATATTTTATTTGCTCAAGATCGTCAATTTCTCCAGACTTAT TTTTCATCACTTAAAAAGGAAGGCTAAGAGTGTAGTTCTGGTGATAGTGT TGGGTCCCTTGGTATTTTTGGTTTGTCACCTTGTGATGAAACACACGTAT ATAAATGTGTGGACAAAAGAATATGAAGGAAATGTGACTTGGAAGATCAA

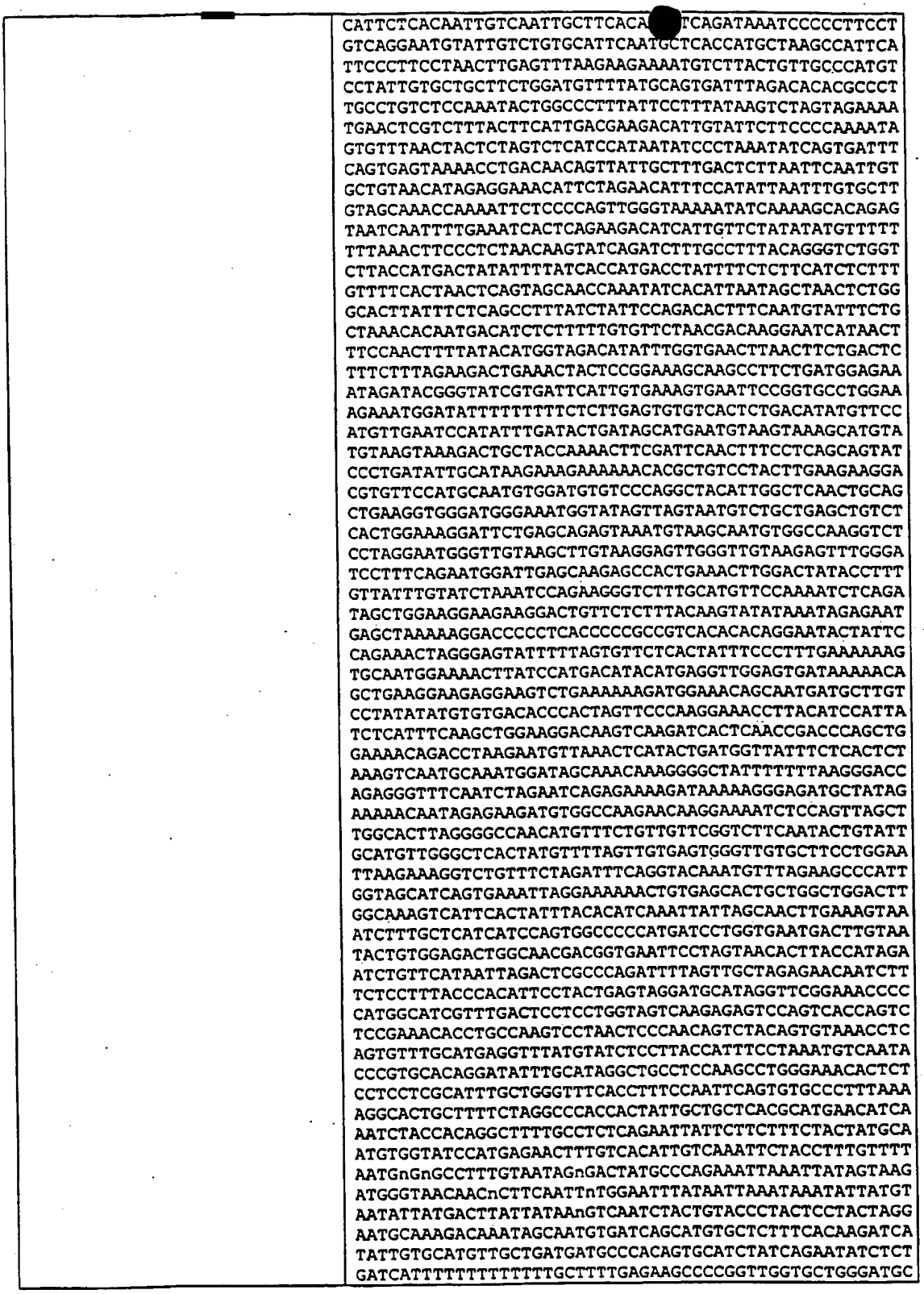
KALQTVTSFLLLCA[redacted]LSMIISVC NFGRLEKQPVFMFCQAIIFSYPSTH PFILILGNKKLKQIFLSVFWQMRYW VKGEKPSSP	ACTGAGGAATGCAATACACCTTTCAAAG[redacted]GACTGTAAGCACACTAGCAA ACTTGATACCCTTCACTCTGACCCTGA[redacted]CTTTTCTGCTGTTAATCTAC TCTCTGTGTAAACATCTCAAGAAGATGCAGCTCCATGGCAAAGGATCTCA AGATCCCAGCACCAAGGTCCACATAAAAGCTTTGCAAACTGTGACCTCCT TTCTTCTGTTATGTGCCATTTACTTTCTGTCCATGATCATATCAGTTTGT AATTTTGGGAGGCTGGAAAAGCAACCTGTCTTCATGTTCTGCCAAGCTAT TATATTCAGCTATCCTTCAACCCACCCATTCATCCTGATTTTGGGAAACA AGAAGCTAAAGCAGATTTTTCTTTCAGTTTTTTGGCAAATGAGGTACTGG GTGAAAGGAGAGAAGCCTTCATCTCCATAG
>hGR37 aa MITFLPIIFSILIVVTFVIGNFANG FIALVNSIEWVKRQKISSADQISHC SGGVQNWFTLGHIITLVCNCV*FGF I*IRSKNFWF*CLSNNQAFQHVGVT SLSIFHLLKTANFSNLIFLHLKKRI KSVGLVILLGPLLFFICNLFVINMD ESVWTKEYEGNVTWKIKLRSAMYHS NMTLTMLANFVPFTLTLISFLLLIC SLCKHLKKMQLHGKGSQDPSTKVHI KALQTVTSFLLLCAIYFLSMIISVC NLGRLEKQPVFMFCEAIIFSYPSTH PFILILGNKKLKQIFLSVLRHVRYW VKGEKPSSS	>hGR37 nt ATGATAACTTTTCTGCCCATCATTTTTTCCATTCTAATAGTGGTTACATT TGTGATTGGAAATTTTGCTAATGGCTTCATAGCTCTAGTAAATTCCATTG AGTGGGTTAAGAGACAAAAGATCTCATCAGCTGACCAAATTTCTCACTGC TCTGGTGGTGTCCAGAATTGGTTTACTCTGGGTCATATTATTACATTGGT ATGCAACTGTGTTTAATTTGGCTTCATATAGATTAGAAGTAAGAATTTTT GGTTCTAATGTCTCAGCAATAACCAAGCATTTCAGCATGTGGGTGTTACT AGCCTCAGCATATTTCATTTGCTCAAGACTGCCAATTTCTCCAACCTTAT TTTTCTCCACCTAAAGAAGAGGATTAAGAGTGTTGGTTTGGTGATACTAT TGGGGCCTTTGCTATTTTTCATTTGTAATCTTTTTGTGATAAACATGGAT GAGAGTGTATGGACAAAAGAATATGAAGGAAACGTGACTTGGAAGATCAA ATTGAGGAGTGCAATGTACCATTCAAATATGACTCTAACCATGCTAGCAA ACTTTGTACCCTTCACTCTGACCCTGATATCTTTTCTGCTGTTAATCTGT TCTCTGTGTAAACATCTCAAGAAGATGCAGCTCCATGGCAAAGGATCTCA AGATCCCAGCACCAAGGTCCACATAAAAGCTTTGCAAACTGTGACCTCCT TTCTTCTGTTATGTGCCATTTACTTTCTGTCCATGATCATATCAGTTTGT AATTTGGGGAGGCTGGAAAAGCAACCTGTCTTCATGTTCTGCGAAGCTAT TATATTCAGCTATCCTTCAACCCACCCATTCATCCTGATTTTGGGAAACA AGAAGCTAAAGCAGATTTTTCTTTCAGTTTTGCGGCATGTGAGGTACTGG GTGAAAGGAGAGAAGCCTTCATCTTCATAG
>hGR38 aa MLTLTRIRTVSYEVRSTFLFISVLE FAVGFLTNAFVFLVNFWDVVKRQPL SNSDCVLLCLSISRLFLHGLLFLSA IQLTHFQKLSEPLNHSYQAIIMLWM IANQANLWLAACLSLLYCSKLIRFS HTFLICLASWSPGRSPVPS	>hGR38 nt
>hGR39 aa LRNAGLNDSNAKLVRNNDLLLINLI LLLPLSVFVMCTSMLFVSLYKHMHW MQSESHKLSSARTEAHINALKTVTT FFCFFVSYFAAFMANMTFRIPYRSH QFFVVKEIMAAYPAGHSVIIVLSNS KFKDLFRRMICLQKE	>hGR39 nt
>hGR40 aa SQYSLGHSYVVIFGYGQMKKTFLGI LWHLKCGLKGRALLATQVGLREKST RSLGVIFLASSYSFFVYVLCH	>hGR40 nt
>hGR41 aa MITFLLIILSILVVFAFVLGNFSNG FIALVNVIDWVNTRKISSADQILTA LAVSRVGLLWVILLHWYANVLNPAL YSSEVIIFISNISAIINHFSIWLAT SLSIFYLLKIVNFSRLIFHHLKRKA KSVVLVIVLGPLVFLVCHLVMKHTY INVWTKEYEGNVTWKIKLRNAIHLS NLTVSTLANLIPFTLTLISFLLLIC SLCKHLKKMQLHSKGSQDPSTKVHI KALQTVTSFLMLFAIYFLYLITSTW NL*TQQSKLVFMFCQTLGIMYPSFH	>hGR41 nt

SFILIMGSRKLKQTF[redacted]LCQVTCL VKGQQPSTP	
>hGR42 aa FIGLTDCIAWMRNQKLCMVGFILTR MALARINIL	
>hGR43 aa LELIFS*KVVATRGLVLGMLGNGLI GLVNCIEWAKSWKVSSADFILTSLA IVRIIRLYLILFDSFIMVLSPHLYT XXXXXXXXXXXXXXXXXXXXXXSL SIFHWFKTANFSNLIFLPLKEED*N VWLGDAVGALGIFHL*SCSENHG*E VCGQKNMKEFCSGMIKLRNAIQLSN LTVTMPANVTPCTLTLISFLLLIYS PCKHVKKMQLHGKGSQHLSTKVHIK VLQTVISFFLLCAIYFVSVIISVWS FKNLENKPVFMFCQAIGFSCSSAHP FILTMGNKKLKQTYLSVLWQMR	
>hGR44 aa MITFLPIIFSILIVVIFVIGNFANG FIALVNSIEWVKRQKISFVDQILTA LAVSRVGLLWVLLLHWYATQLNPAF YSVEVRITAYNVWAVTNHFSSWLAT SLSMFYLLRIANFSNLIFLRIKRRV KSVVLVILLGPLLFLVCHLFVINMD ETVWTKEYEGNVTWKIKLRSAMYHS NMTLTMLANFVPLTLTLISFLLLIC SLCKHLKKMQLHGKGSQDPSTKVHI KALQTVTSFLLLCAIYFLSMIISVC NLGRLEKQPVFMFCQAIIFSYPSTH PFILILGNKKLKQIFLSVLRHVRYW VKDRSLRLHRFTRGALCVF	
>hGR45 aa MATELDKIFLILAIAEFIISMLGNV FIGLVNCSEGIKNQKVFSADFILTC LAISTIGQLLVILFDSFLVGLASHL YTTYRLGKTVIMLWHMTNHLTTWLA TCLSIFYFFKIAHFPHSLFLWLRWR MNGMIVMLLILSLFLLIFDSLVLEI FIDISLNIIDKSNLTLYLDESKTLY DKLSILKTLLSLTSFIPFSLFLTSL LFLFLSLVRHTRNLKLSSLGSRDSS TEAHRRAMKMVMSFLFLFIVHFFSL QVANWIFFMLWNNKCIKFVMLALNA FPSCHSFILILGNSKLQQTAVRLLW HLRNYTKTPNPLPL	
>hGR46 MSFLHIVFSILVVVAFILGNFANGF IALINFIAWVKKQKISSADQIIADK QSPELVCSG	
>hGR47 aa MLNALYSILIIIINI*FLIGILGNG FITLVNGIDWVKM*KRSSILTALTI SRICLISVIMVRWFI	
>hGR48 aa VSRVGLLWVILLHWYSTVLNPTSSN	

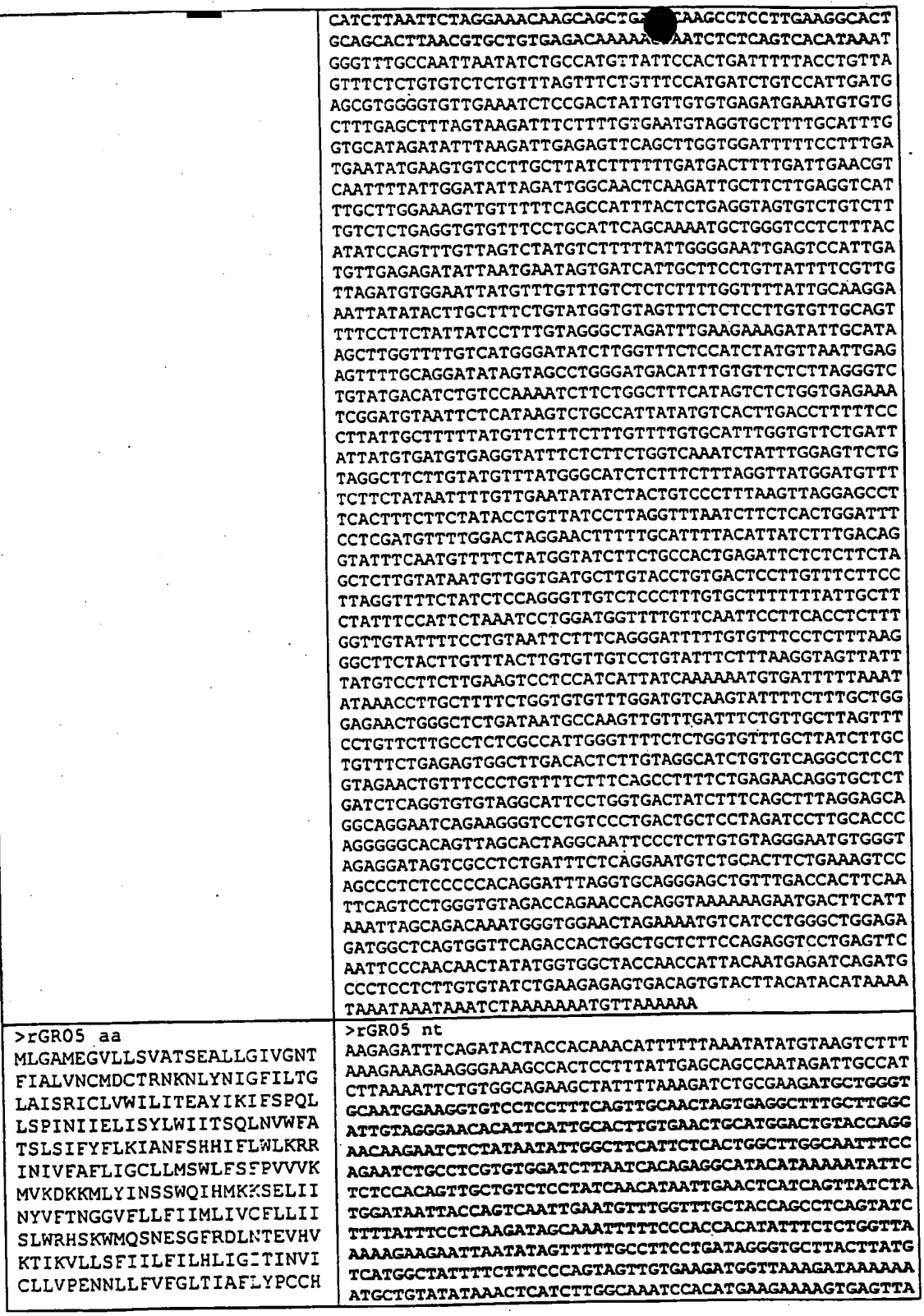
LKVIIFISNAWAVTNF IWLATSL SIFYLLKIVN	
>hGR49 aa TVTMLANLVPFTVTLISFLLLVCSL CKHLKKMHLHGKGSQDPSTKVHIKV LQTVISFLLLCAIYFVSVIISS	
>hGR50 aa MITFLPIIFSILVVVTFVIGNFANG FIALVNSTEWVKRQKISFADQIVTA LAVSRVGLLWVLLLNWYSTVLNPAF YSVELRTTAYNIWAVTGHFSNWPAT SLSIFYLLKIANFSNLIFLRLKRRV KSVILVVLLGPLLFLACHLFVVNMN QIVWTKEYEGNMTWKIKLRRAMYLS DTTVTMLANLVPFTVTLISFLLLVC SLCKHLKKMQLHGKGSQDPSTKVHI KVLQTVISFFLLCAIYFVSVIISVW SFKNLENKPVFMFCQAIGFSCSSAH PFILIWGNKKLKQTYLSVLWQMRY	

>rGR01 aa MMEGHILFFFLVVMVQFVTGVLANG LIVVVHAIDLIMWKKMAPLDLLLFC LATSRIILQLCILFAQLCLFSLVRH TLFEDNITFVFIINELSLWFATWLG VFYCAKIATIPHPLFLWLKMRISRL VPWLILGSVLYVIITTFIHSRETSA ILKPIFISLFPKNATQVGTGHATLL SVLVLGLTLPLFIFTVAVLLLIYSL WNYSRQMRTMVGTREYSGHAHISAM LSILSFLILYLSHYMVAVLISTQVL YLGSRTFVFCLLVIGMYPSIHSIVL ILGNPKLKRNAKMFIVHCKCCHCTR AWVTSRSPRLSDLPVPPTHPSANKT SCSEACIMPS	>rGR01 nt CAGGAATCATAAATGGCTGAAACTGGGCAGAACTCTATGCATTATTTAAAG AAGTCATTGGTTTGTCATTCTTAAAATGATGGAAGGGCATATACTCTTCTT CTTTTTGGTTGTGATGGTGCAGTTTGTCACTGGGGTCTTGGCAAATGGCCT CATTGTGGTTGTCCATGCTATTGACTTGATCATGTGGAAGAAAATGGCCCC GTTGGATCTGCTTCTATTTTGCCTGGCGACTTCTCGGATCATTCTGCAGTT ATGTATATTGTTTGCACAATTGTGTCTATTCTCTTTGGTGAGACACACTTT ATTTGAGGACAATATTACCTTTGTCTTCATCATAAATGAACTGAGTCTTTG GTTTGCTACATGGCTCGGTGTTTTCTACTGTGCCAAGATTGCTACCATTCC TCACCCACTCTTTCTGTGGCTGAAGATGAGGATATCCAGGTTGGTACCATG GCTGATCCTGGGATCTGTGCTCTATGTAATTATTACTACTTTCATCCATAG CAGAGAGACTTCAGCAATCCTTAAACCAATTTTTATAAGCCTTTTTCCTAA AAATGCAACTCAAGTCGGAACAGGGCATGCCACACTACTCTCAGTCCTGGT CCTTGGGCTCACACTGCCGTTGTTCATCTTTACTGTTGCTGTTCTGCTCTT GATATACTCCCTGTGGAATTATAGCAGGCAGATGAGGACTATGGTAGGCAC CAGGGAGTATAGCGGACATGCTCACATCAGTGCAATGCTGTCCATTCTATC ATTCCTCATCCTCTATCTCTCCCACTACATGGTGGCTGTTCTGATCTCTAC TCAAGTCCTCTACCTTGGAAGCAGAACCTTTGTATTCTGCTTACTGGTTAT TGGTATGTACCCCTCAATACACTCGATTGTCTTAATTTTAGGAAATCCTAA GCTGAAACGAAATGCAAAAATGTTCATTGTCCATTGTAAGTGTTGTCATTG TACAAGAGCTTGGGTCACCTCAAGGAGCCCAAGACTCAGTGACTTGCCAGT GCCTCCTACTCATCCCTCAGCCAACAAGACATCCTGCTCAGAAGCCTGTAT AATGCCATCCTAATTGTCCAGCCTGAGGTTTAATCCTAGGTTTGGTACTAT TTCAAAGAGTAAAGTTGATCATTAAAGCACAACATATGTTGGTGGATGACA TCAAGGTCCATATCCCAGTTGTCAATTGTAAACCTCACCTTGCAAGATGAT GTCACTGAGAAAGCAGGACAAATGGAGTCTAGGTCCTTCTGTATGACTTGC TGCAGTATATGTGAATCTATAATTTTCTCCAAAAAAACAAAAAAAAAAAAA AAAAA
>rGR02 aa MFSQKTNYSHLFTFSIIFYVEIVTG ILGNGFIALVNIMDWLKRRRISTAD QILTALALTRLIYVWSVLICILLLF LCPHLSMRPEMFTAIGVIWVVDNHF SIWLATCLGVFYFLKIASFSNSLFL YLKWRVKKVVLMIILISLIFLMLNI SSLGMYDHFSIDVYEGNMSYNLVDS THFPRIFLFTNSSKVFLIANSSHVF	>rGR02 nt (3'UTR not pristine) ATTTTGCTCCACTATTTTGCTCTTCTGCAGTAACACAGACCACAAAACAAT GGAGCCAATGGGTCAAGAGCTGAAACTTCAGGAAGTGGGAGCCAAATTTTC TTTGTGATAGGTTGGCATATGAGAATTCATTATTTGATGCAGCTTCTGAAA ACTGGATGTGAAATACTGGATGAAGCAGAGGTGATGACCCCTTTGAAATTA AAAAGCCAAGATGTTCATGGAGAAATTATAAAACAATATCTGGGAAATTTG ATGCTTCCTAATCGGGTGTAAATGGGATTTTAAATGATGAACATTTTGAAT TTCCAATGACCATTATGTAAAGTTTTTAAACACAGTAGAGACATCATAAAT TGAAGCATGTTCTCACAGAAAACAAACTACAGCCATTTGTTTACTTTTTCA ATTATTTTTTATGTGGAAATAGTAACAGGAATCTTAGGAAATGGATTCATA

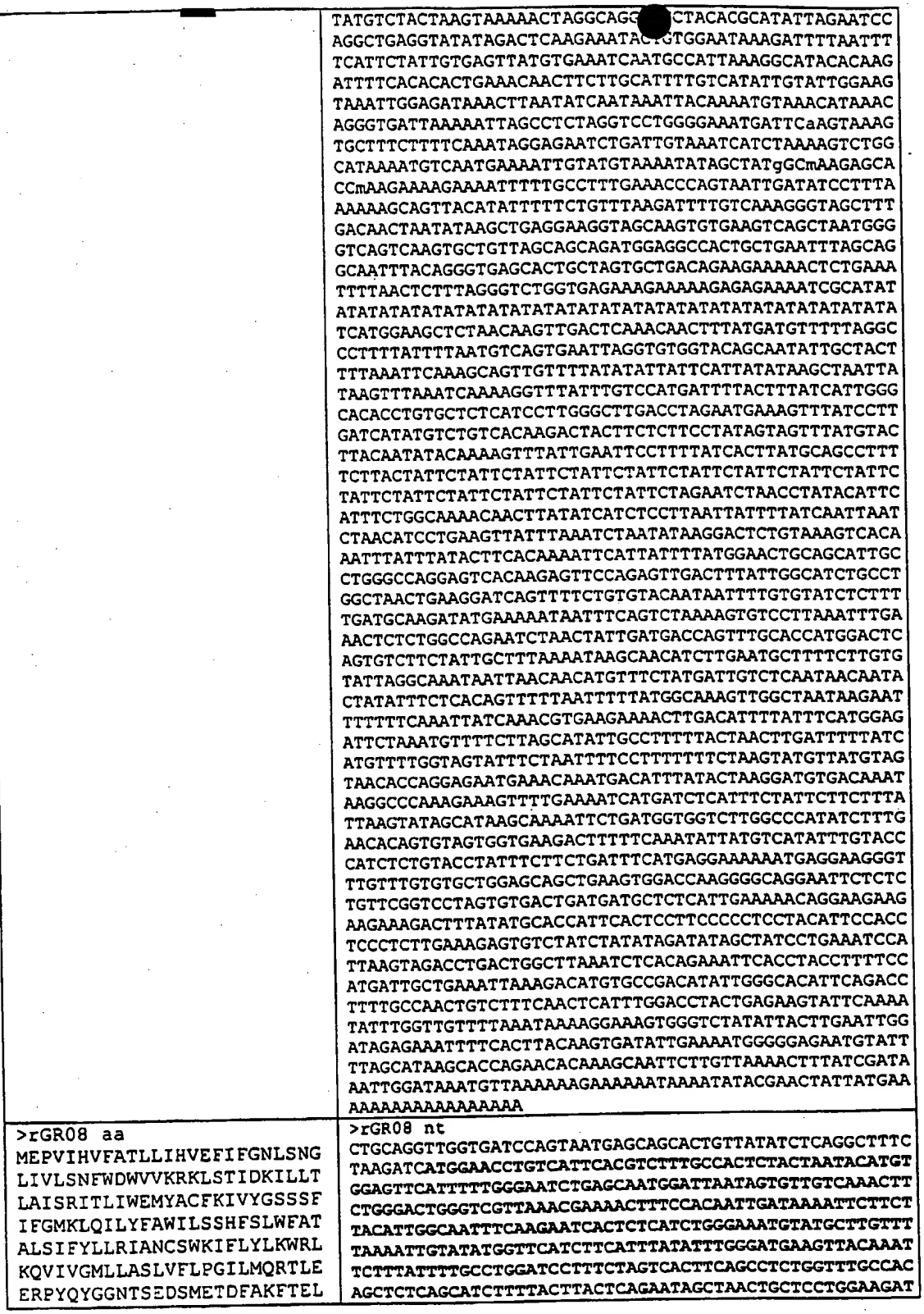
LPINSLFMLIPFTVSL[redacted]FVLFLS LWKHHKKMQVNAKGPRDASTMAHTK ALQIGFSFLLLYAIYLLFIITGILN LDLMRCIVILLFDHISGAVFSISHS FVLILGNSKLRQATLSVLPCLRCRS KDMDTVVF	GCACTAGTGAATATCATGGACTGGCTCAA[redacted]GAGGAGGATCTCTACTGCA GATCAGATTCTCACTGCTTTGGCCCTTAC[redacted]GACTCATTTATGTGTGGTCT GTACTCATTTGTATATTGTTACTATTTCTGTGCCCACATTTGTCTATGAGA CCAGAAATGTTTACAGCGATAGGTGTTATCTGGGTAGTGGATAACCACTTC AGCATCTGGCTTGCTACATGTCTTGGTGTCTTTTATTTCCTCAAAATAGCC AGTTTTTCTAACTCTTTGTTTCTTTACCTAAAGTGGAGAGTTAAAAAAGTG GTTTTAATGATAATACTGATATCACTGATTTTCTTGATGTTAAACATTTCA TCATTAGGGATGTATGATCATTTCTCAATTGATGTTTATGAAGGTAATATG TCTTATAATTTGGTGGATTCAACACATTTTCCCAGAATTTTCTTATTCACA AACTCATCTAAGGTCTTCTTAATCGCCAATTCATCCCATGTTTTCTTACCC ATCAACTCACTCTTCATGCTCATACCCTTCACAGTTTCCCTGGTAGCTTTT TTCGTGCTCTTTCTCTCACTGTGGAAGCATCACAAGAAGATGCAGGTCAAT GCCAAAGGACCCAGAGATGCCAGCACCATGGCCCACACAAAAGCCTTGCAA ATTGGGTTCTCCTTCCTCCTGCTGTATGCAATATACTTACTTTTCATTATC ACAGGAATTTTGAACCTTGACTTGATGAGATGTATAGTAATACTTTTATTT GACCACATATCTGGAGCAGTTTTTTCTATAAGCCACTCATTTGTGCTGATT CTGGGAAACAGTAAGCTGAGACAAGCCACTCTTTCTGTGCTGCCTTGTCTT AGGTGCCGGTCCAAAGATATGGACACTGTCGTTTTCTAATAAATTCCAGAG TACATTATGCAAAATCTTGAGGGTGATCAGTTCATAGAAAAAGTAATCTTA GAGGGGAAAATAAAATATTGGGGCTTCAAATGTTGGATGGGTAATACATAG GAAGGCAGGACAAGGATGAAGGAGACTAGCATTATATAAGTGATTTCACAG GGGAAATGGGAAAGAGGGCTTTTATATAATGAAGAAGAAGATAAATGATGA AGGATGAGGAAGAGTTAAATATGTAAAATGACAATAGAGATGGCATCATGC CGTTTTAAGAAATTTGGAATGCATATGTATGTTTATATATTTTTTAATTTT TATTGAATATATTTATTTACATTTTAAATGTTATCCTGTTTCCCCCACCCA ACCTCCCACCTCTTCCCACCTCCTTGCCCTGACATTCCCCTGCACTGGGGA ATCCAGCCTTGACAGGACCAAGGGCTTCTCCTCCCTTTGTTGCCAACAAGG CCATTCTTTGCTACATGTGCAGCAGGAGCCATGGATCTGTCTATGTGTACT CTTTGGATGGTGGTTTAGTCCCTGGGAGCTCTTGTTGGTTGGTATTGTTGT TCTTATGGTGTTGCAACTCCCTTCAGCTCCTTCAATCCTTCCTGTAACTCC TCCAATGTGGACCCTGTTCTCAGTCCAATGGTTGACTATGAGCATTCACCT CTGTGATTGTCATGCTCTGGCACAGCTTCTCAGAAGACAGCTACATCAGTC TCCTATAAGAGTGCACTTCATGGCATCAGCAATGTTGTCTTGATTTGGTGT CTGTATGTATATGGGCTGGATCCCAGGTGGGGCAGGCGCTGAATGGTCATT CCTTCAGTCTTTGCTCCAAACTTTGTCTTTATATCTCCTATGAATATTTTT GTTCCCCCTTATAAGAATGACTGAAGTATCCACACTTTGGCCATCCTTCTT CATGAGCTTCATGTGGTCTGTGAATTGTACATTGTGTAATCCAAGCTTTTG GGCTAATATCCAATTATAGTGAGTGCATACCAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
>rGR03 aa MVPTQVTIFSIIMYVLESLVIIVQS CTTVAVLFREWMHFQRLSPVEIILI SLGISHFCLQWTSMLYNFGTYSRPV LLFWKVSVVWEFMNVLTFWLTSLLA VLYCVKVSSFSHPVFLWLRLKILKL VLWLLLGALIASCLSIIPSVVKYHI QMELLTLDHLPKNSSLILRLQMFEW YFSNPFKMIGFGVPFLVFLISIILL TVSLVQHWGQMKHYSSSSSSLRAQC TVLKSLATFFIFFTSYFLTIVVSFI GTVFDKKSWFWVCEAVIYGLVCIHF TSLMMSNPTLKKALRLQFWSPESS	>rGR03 nt (cds pristine; 3'UTR not so hot) GCATGGTGCCAACCCAAGTCACCATCTTCTCTATCATCATGTATGTGCTTG AGTCCTTAGTCATAATTGTGCAAAGTTGCACAACGGTTGCAGTGCTGTTCA GAGAGTGGATGCACTTTCAAAGACTGTCGCCGGTGGAAATAATTCTCATCA GCCTGGGCATTTCACATTTCTGTCTACAGTGGACATCGATGCTGTACAACT TTGGTACCTACTCTAGGCCTGTCCTTTTATTTTGGAAGGTATCGGTCGTCT GGGAGTTCATGAACGTTTTGACATTCTGGCTAACCAGTTTGCTTGCTGTCC TCTACTGTGTCAAGGTCTCTTCCTTCTCTCACCCCGTCTTCCTCTGGCTGA GGTTGAAAATTTTGAAACTGGTTCTCTGGTTGCTATTGGGCGCTCTGATAG CTTCTTGTTTGTCAATCATCCCTTCTGTTGTTAAATATCATATCCAGATGG AATTACTCACCCTAGATCATTTACCCAAAAACAGTTCTTTGATTCTAAGAC TGCAAATGTTCGAGTGGTATTTTTCTAATCCTTTCAAAATGATTGGGTTTG GCGTTCCTTTCCTCGTGTTCCTGATTTCTATCATCTTACTCACAGTCTCGC TGGTCCAGCATTGGGGGCAGATGAAACACTACAGCAGCAGCAGCTCCAGCC TGAGAGCTCAGTGCACTGTTCTGAAGTCTCTTGCCACCTTCTTCATCTTCT TCACATCCTATTTTCTGACTATAGTCGTCTCCTTTATTGGCACCGTGTTTG ATAAGAAGTCATGGTTCTGGGTCTGCGAAGCTGTCATCTATGGTTTAGTCT GTATTCACTTCACTTCCCTGATGATGAGCAACCCTACACTGAAAAAAGCAC TCAGGTTGCAGTTCTGGAGCCCAGAGTCTTCCTAAGGCAGGGAATTCAGTG AAGCCTCTGGGGTAAGGAGGCTTTGCATTGGCACAGTTCTTAGAGTGAAAT GCAAACGTGGACACGAACTTCATTCTCTTTCATGTCCACAGATGGATGGAT CTATAAATCATCACCAATCTTCCCTGTATTCTGACCCATCCTTTTCCTGTC CTATCCATAGTCCCCAGGTTGGTTTTGATTTTTCTCATGATCACACCTTAG CTTTAGCCACCGTTGCAATATCAAACATGATCTATATGTTACAGCCAAAAT


CATTCTCACAATTGTCAATTGCTTCACA[redacted]TCAGATAAATCCCCCTTCCT
GTCAGGAATGTATTGTCTGTGCATTCAATGCTCACCATGCTAAGCCATTCA
TTCCCTTCCTAACTTGAGTTTAAGAAGAAAATGTCTTACTGTTGCCCATGT
CCTATTGTGCTGCTTCTGGATGTTTTATGCAGTGATTTAGACACACGCCCT
TGCCTGTCTCCAAATACTGGCCCTTTATTCCTTTATAAGTCTAGTAGAAAA
TGAACTCGTCTTTACTTCATTGACGAAGACATTGTATTCTTCCCCAAAATA
GTGTTTAACTACTCTAGTCTCATCCATAATATCCCTAAATATCAGTGATTT
CAGTGAGTAAAACCTGACAACAGTTATTGCTTTGACTCTTAATTCAATTGT
GCTGTAACATAGAGGAAACATTCTAGAACATTTCCATATTAATTTGTGCTT
GTAGCAAACCAAAATTCTCCCCAGTTGGGTAAAAATATCAAAAGCACAGAG
TAATCAATTTTGAAATCACTCAGAAGACATCATTGTTCTATATATGTTTTT
TTTAAACTTCCCTCTAACAAGTATCAGATCTTTGCCTTTACAGGGTCTGGT
CTTACCATGACTATATTTTATCACCATGACCTATTTTCTCTTCATCTCTTT
GTTTTCACTAACTCAGTAGCAACCAAATATCACATTAATAGCTAACTCTGG
GCACTTATTTCTCAGCCTTTATCTATTCCAGACACTTTCAATGTATTTCTG
CTAAACACAATGACATCTCTTTTTGTGTTCTAACGACAAGGAATCATAACT
TTCCAACTTTTATACATGGTAGACATATTTGGTGAACTTAACTTCTGACTC
TTTCTTTAGAAGACTGAAACTACTCCGGAAAGCAAGCCTTCTGATGGAGAA
ATAGATACGGGTATCGTGATTCATTGTGAAAGTGAATTCCGGTGCCTGGAA
AGAAATGGATATTTTTTTTTCTCTTGAGTGTGTCACTCTGACATATGTTCC
ATGTTGAATCCATATTTGATACTGATAGCATGAATGTAAGTAAAGCATGTA
TGTAAGTAAAGACTGCTACCAAAACTTCGATTCAACTTTCCTCAGCAGTAT
CCCTGATATTGCATAAGAAAGAAAAAACACGCTGTCCTACTTGAAGAAGGA
CGTGTTCCATGCAATGTGGATGTGTCCCAGGCTACATTGGCTCAACTGCAG
CTGAAGGTGGGATGGGAAATGGTATAGTTAGTAATGTCTGCTGAGCTGTCT
CACTGGAAAGGATTCTGAGCAGAGTAAATGTAAGCAATGTGGCCAAGGTCT
CCTAGGAATGGGTTGTAAGCTTGTAAGGAGTTGGGTTGTAAGAGTTTGGGA
TCCTTTCAGAATGGATTGAGCAAGAGCCACTGAAACTTGGACTATACCTTT
GTTATTTGTATCTAAATCCAGAAGGGTCTTTGCATGTTCCAAAATCTCAGA
TAGCTGGAAGGAAGAAGGACTGTTCTCTTTACAAGTATATAAATAGAGAAT
GAGCTAAAAAGGACCCCCTCACCCCCGCCGTCACACACAGGAATACTATTC
CAGAAACTAGGGAGTATTTTTAGTGTTCTCACTATTTCCCTTTGAAAAAAG
TGCAATGGAAAACTTATCCATGACATACATGAGGTTGGAGTGATAAAAACA
GCTGAAGGAAGAGGAAGTCTGAAAAAAGATGGAAACAGCAATGATGCTTGT
CCTATATATGTGTGACACCCACTAGTTCCCAAGGAAACCTTACATCCATTA
TCTCATTTCAAGCTGGAAGGACAAGTCAAGATCACTCAACCGACCCAGCTG
GAAAACAGACCTAAGAATGTTAAACTCATACTGATGGTTATTTCTCACTCT
AAAGTCAATGCAAATGGATAGCAAACAAAGGGGCTATTTTTTTAAGGGACC
AGAGGGTTTCAATCTAGAATCAGAGAAAAGATAAAAAGGGAGATGCTATAG
AAAAACAATAGAGAAGATGTGGCCAAGAACAAGGAAAATCTCCAGTTAGCT
TGGCACTTAGGGGCCAACATGTTTCTGTTGTTCGGTCTTCAATACTGTATT
GCATGTTGGGCTCACTATGTTTTAGTTGTGAGTGGGTTGTGCTTCCTGGAA
TTAAGAAAGGTCTGTTTCTAGATTTCAGGTACAAATGTTTAGAAGCCCATT
GGTAGCATCAGTGAAATTAGGAAAAAACTGTGAGCACTGCTGGCTGGACTT
GGCAAAGTCATTCACTATTTACACATCAAATTATTAGCAACTTGAAAGTAA
ATCTTTGCTCATCATCCAGTGGCCCCCATGATCCTGGTGAATGACTTGTAA
TACTGTGGAGACTGGCAACGACGGTGAATTCCTAGTAACACTTACCATAGA
ATCTGTTCATAATTAGACTCGCCCAGATTTTAGTTGCTAGAGAACAATCTT
TCTCCTTTACCCACATTCCTACTGAGTAGGATGCATAGGTTCGGAAACCCC
CATGGCATCGTTTGACTCCTCCTGGTAGTCAAGAGAGTCCAGTCACCAGTC
TCCGAAACACCTGCCAAGTCCTAACTCCCAACAGTCTACAGTGTAAACCTC
AGTGTTTGCATGAGGTTTATGTATCTCCTTACCATTTCCTAAATGTCAATA
CCCGTGCACAGGATATTTGCATAGGCTGCCTCCAAGCCTGGGAAACACTCT
CCTCCTCGCATTTGCTGGGTTTCACCTTTCCAATTCAGTGTGCCCTTTAAA
AGGCACTGCTTTTCTAGGCCCACCACTATTGCTGCTCACGCATGAACATCA
AATCTACCACAGGCTTTTGCCTCTCAGAATTATTCTTCTTTCTACTATGCA
ATGTGGTATCCATGAGAACTTTGTCACATTGTCAAATTCTACCTTTGTTTT
AATGnGnGCCTTTGTAATAGnGACTATGCCCAGAAATTAAATTATAGTAAG
ATGGGTAACAACnCTTCAATTnTGGAATTTATAATTAAATAAATATTATGT
AATATTATGACTTATTATAAnGTCAATCTACTGTACCCTACTCCTACTAGG
AATGCAAAGACAAATAGCAATGTGATCAGCATGTGCTCTTTCACAAGATCA
TATTGTGCATGTTGCTGATGATGCCCACAGTGCATCTATCAGAATATCTCT
GATCATTTTTTTTTTTTTGCTTTTGAGAAGCCCCGGTTGGTGCTGGGATGC

	TTCATAGCAGGTCCACCATAGACACATG[redacted]AGAGGAAAGCTGCCTCTCTC TCTTCATTCCCAAGGAACAGTAAAAGCAGA[redacted]AAGGCTCTTATGTTCTAAAG AACAGAAAATAGCCTGCATTTCAACTACCTCCTGTTCAGAAGGCACCGAAA CACACCACCAAGCAAGACACCCCTTTACTTTCTCCTGCTTCCCTCAATTTG ATGATCATTTGGAAATAAGAAGAAAGAAAAAGATGTGGAAGCCAATTAAAA ACAGTCTTGTCTATCTCCCTGGTGAGCTCTCAACTTCTTAGTCAGACCAAA GTAGGTGAAAAAATAATAATTTTTAATTTGGTATGAGAGTCATGTTTAGGC TGAAAATCTTAAAAAATCTTAGCATAAAAACATTTTCCCCTAGACCCATGA AATTTATAATATTATCTGTGGTTGAGAAAGGCTAGTTATAGAAAAATGTTT AGAATCAGAATATTTTGAGGGCTCTTTTTTTGTTTTGCCTAATCATTACAT TTGTTATAAGAAGTCTAAAAGTTGGTATGCTACAGGTCTTGTCATATTTTC TCTGAGGTTGAGTGCCAAGTAGTCTGCATTGTGTTTAAATCCTGCTTAAAA TTATCCCAAGACAATATAACTTCTCAGGAGCTAAGCCAAGGGCCCCTTTCA GACTACCTTAGTCCTCTCTCACCGTTGTCACCGTGGCTCATACATCAGAAT CCTGAGGGAGCATCATGAAATCTAAGGCTTTACAACAGAATCTTTCTATCC CTGGTAGAAATCTTTTAACCTTGGGTTTTATTCTCATGCCATTCTGATGCT CGTATTTAAATTTTATGTGTTTTTTCATATGTTCTTGCATTTCTATCGTTA AATTATGGTGACATACTTTCAAATGCTTTGTTATTTTAAAAAGGGACAAAG AGAGATAGAAAGACAGGGAAAGATAGACAGAGGCTTGCCTAATACAGTCAA GAAAGAAGCTATCAAAAGTATTTAGCAATACAACATTTATGATATATTCAT AACTGTTAACCATTTTTAATATTCTAAAATTTCACTTTTGTTTCAGAAATG TATATTAAGAGAATCTGAGAAACATTTTTTTCTCATAGATGTAGAAAAACA CACAAAATAAGGTATAACACATTTAAGTGATTGAAAATAAAAACAAAAGCT TGCAAACAGGAGGAAAAGTACATTGTAGGCTTTCGACATGGAGCTGCTACT AGGACCCAGGACTTGTTTATCATTTATTTGCCAAGTCCCACAAACTCAGGG CAATACATCTCTGAGACAGTTTCCTATATTTTAATAAAACTTCCAAAATTG ATACTCAGTGTGAATTGGCTAGCTTTAATGGCAGTCATTGGATAAACAATT CCAATGCCAAATTTCCCTAAGTTGATATATTTGATTAATATGTATATTAAA ACATCAGGCTATCCATCGGTTGGATCAAATACATTCTTTAGGGATCCATTC TTTTCCTTAAATTTGACTTATATGTGGATTCTTTTCACAATAAATAAGTAA ATGAGCATTTATTTTAAAACTATTTTAGACGGAACTGAATTACAGCCAAGG TAGTCAAAATGACTGAGAATAATCACTTACATATTTACAAGGGAAAGTGAC TCTTCAGATTTAAGTTTAAAATTAGAAGAGAGATAAATTTCACAAGCTTTC ACTCCTAAGGCTAAAGATAGGCTGTGTAGGTAGTTATTTCTGAGCACATTG GCACATCACCATTGTCAGTACTTGAGGGTTTGAATGAAGCTCACTCAAAGA ACTTGGAAAGAAGGTGGTCTTCTGACATCAATCAAGAAACAAGCTTTCCTC CCTACTTCTTCCCTAAATGCAACAACCTAAGAATTATCCACAAGATGGATG GCGCAAGGGTTCCTCAATCAATTTCAGGATGTACATCAATGCGCAGCCTAT ACTACACCGAAAAGGAAGCGCATGGGTCTTAAAAAGTAAAGGGGATATCAA AAAATTCGCAACCAAACAAAAAGTGGCACACATTTAAGCTAGGTCTATGTT TGGTCAGTTACACCTGGAGAAGGGGGACATTTGGTCAGCTCATTCGAACAC TGTCAAGTCCTACCAACAATTCCTCTATGCTATTACCCATTAAACCTCAGG TCTCATCGAAAAAAAAAAAAAAAAA
>rGR04 aa MLSAAEGILLCVVTSEAVLGVLGDT FIALANCMEYAKNKKLSKIGFILIG LAISRIGVVWIIILQGYMQVFFPHI LTFGNITEYITYIWVFLNHLSVWFA TNLNILYFLKIANFSNSVFLWLKSR VRVVFIFLSGCLLTSWLLCFPQFSK MLNNSKMYWGNTSWLQQQKNVFLIN QSLTNLGIFFFIIVSLITCFLLIVF LWRHIRQMHSDGSGLRDLNTEAHVK AMRVLISFAVLFILHFVGLSIQVLC FFLPQNNLLFITGLIATCLYPCGHS IILILGNKQLKQASLKALQHLTCCE TKRNLSVT	>rGR04 nt (pristine cds; 3'UTR not so hot) TGGTTCCATCACATGACAATAGGCTTGAAAAACTTGCAGATAGAGAAGACA TAACCCCTCCAACAAGAAGCCAACATATGGGACATTCTCCAGCAGATAATT TATAACAGATGCAACGGGAGCAACTTCGAGATCTGCAAAGATGCTGAGTGC AGCAGAAGGCATCCTCCTTTGTGTTGTCACTAGTGAGGCAGTGCTGGGGGT TTTAGGAGACACATTCATTGCACTTGCAAACTGCATGGAGTATGCCAAGAA CAAGAAGCTCTCTAAGATTGGTTTCATTCTCATTGGCTTGGCGATTTCCAG AATTGGTGTCGTATGGATAATAATTTTACAGGGGTATATGCAAGTATTTTT TCCACACATACTTACCTTTGGAAACATAACTGAATATATTACTTACATATG GGTGTTTCTCAATCACTTAAGTGTCTGGTTTGCTACCAACCTCAATATCCT CTACTTTCTAAAGATAGCAAATTTTTCCAACTCTGTATTTCTCTGGCTGAA AAGTAGAGTCCGTGTGGTTTTTATCTTTCTGTCAGGATGCTTACTTACCTC GTGGTTACTATGTTTTCCACAATTTTCAAAGATGCTTAACAACAGTAAAAT GTACTGGGGAAACACGTCTTGGCTCCAGCAGCAGAAAAATGTCTTCCTTAT TAACCAAAGTTTAACCAATCTGGGAATCTTCTTTTTCATTATTGTATCCCT GATTACCTGCTTCCTGTTGATTGTTTTCCTCTGGAGACACATCAGGCAAAT GCACTCAGATGGTTCAGGACTCAGAGACCTCAACACAGAAGCTCATGTGAA AGCCATGAGAGTTCTAATATCTTTTGCGGTACTCTTTATCCTGCATTTCGT AGGTCTTTCCATACAAGTGCTATGCTTTTTTCTGCCACAAAACAACCTACT CTTTATAACTGGTTTGATAGCCACATGCCTCTATCCCTGTGGTCACTCAAT


CATCTTAATTCTAGGAAACAAGCAGCTG[redacted]AAGCCTCCTTGAAGGCACT
GCAGCACTTAACGTGCTGTGAGACAAAAA[redacted]AATCTCTCAGTCACATAAAT
GGGTTTGCCAATTAATATCTGCCATGTTATTCCACTGATTTTTACCTGTTA
GTTTCTCTGTGTCTCTGTTTAGTTTCTGTTTCCATGATCTGTCCATTGATG
AGCGTGGGGTGTTGAAATCTCCGACTATTGTTGTGTGAGATGAAATGTGTG
CTTTGAGCTTTAGTAAGATTTCTTTTGTGAATGTAGGTGCTTTTGCATTTG
GTGCATAGATATTTAAGATTGAGAGTTCAGCTTGGTGGATTTTTCCTTTGA
TGAATATGAAGTGTCCTTGCTTATCTTTTTTGATGACTTTTGATTGAACGT
CAATTTTATTGGATATTAGATTGGCAACTCAAGATTGCTTCTTGAGGTCAT
TTGCTTGGAAAGTTGTTTTTCAGCCATTTACTCTGAGGTAGTGTCTGTCTT
TGTCTCTGAGGTGTGTTTCCTGCATTCAGCAAAATGCTGGGTCCTCTTTAC
ATATCCAGTTTGTTAGTCTATGTCTTTTTATTGGGGAATTGAGTCCATTGA
TGTTGAGAGATATTAATGAATAGTGATCATTGCTTCCTGTTATTTTCGTTG
TTAGATGTGGAATTATGTTTGTTTGTCTCTCTTTTGGTTTTATTGCAAGGA
AATTATATACTTGCTTTCTGTATGGTGTAGTTTCTCTCCTTGTGTTGCAGT
TTTCCTTCTATTATCCTTTGTAGGGCTAGATTTGAAGAAAGATATTGCATA
AGCTTGGTTTTGTCATGGGATATCTTGGTTTCTCCATCTATGTTAATTGAG
AGTTTTGCAGGATATAGTAGCCTGGGATGACATTTGTGTTCTCTTAGGGTC
TGTATGACATCTGTCCAAAATCTTCTGGCTTTCATAGTCTCTGGTGAGAAA
TCGGATGTAATTCTCATAAGTCTGCCATTATATGTCACTTGACCTTTTTCC
CTTATTGCTTTTTATGTTCTTTCTTTGTTTTGTGCATTTGGTGTTCTGATT
ATTATGTGATGTGAGGTATTTCTCTTCTGGTCAAATCTATTTGGAGTTCTG
TAGGCTTCTTGTATGTTTATGGGCATCTCTTTCTTTAGGTTATGGATGTTT
TCTTCTATAATTTTGTTGAATATATCTACTGTCCCTTTAAGTTAGGAGCCT
TCACTTTCTTCTATACCTGTTATCCTTAGGTTTAATCTTCTCACTGGATTT
CCTCGATGTTTTGGACTAGGAACTTTTTGCATTTTACATTATCTTTGACAG
GTATTTCAATGTTTTCTATGGTATCTTCTGCCACTGAGATTCTCTCTTCTA
GCTCTTGTATAATGTTGGTGATGCTTGTACCTGTGACTCCTTGTTTCTTCC
TTAGGTTTTCTATCTCCAGGGTTGTCTCCCTTTGTGCTTTTTTTATTGCTT
CTATTTCCATTCTAAATCCTGGATGGTTTTGTTCAATTCCTTCACCTCTTT
GGTTGTATTTTCCTGTAATTCTTTCAGGGATTTTTGTGTTTCCTCTTTAAG
GGCTTCTACTTGTTTACTTGTGTTGTCCTGTATTTCTTTAAGGTAGTTATT
TATGTCCTTCTTGAAGTCCTCCATCATTATCAAAAAATGTGATTTTTAAAT
ATAAACCTTGCTTTTCTGGTGTGTTTGGATGTCAAGTATTTTCTTTGCTGG
GAGAACTGGGCTCTGATAATGCCAAGTTGTTTGATTTCTGTTGCTTAGTTT
CCTGTTCTTGCCTCTCGCCATTGGGTTTTCTCTGGTGTTTGCTTATCTTGC
TGTTTCTGAGAGTGGCTTGACACTCTTGTAGGCATCTGTGTCAGGCCTCCT
GTAGAACTGTTTCCCTGTTTTCTTTCAGCCTTTTCTGAGAACAGGTGCTCT
GATCTCAGGTGTGTAGGCATTCCTGGTGACTATCTTTCAGCTTTAGGAGCA
GGCAGGAATCAGAAGGGTCCTGTCCCTGACTGCTCCTAGATCCTTGCACCC
AGGGGGCACAGTTAGCACTAGGCAATTCCCTCTTGTGTAGGGAATGTGGGT
AGAGGATAGTCGCCTCTGATTTCTCAGGAATGTCTGCACTTCTGAAAGTCC
AGCCCTCTCCCCCACAGGATTTAGGTGCAGGGAGCTGTTTGACCACTTCAA
TTCAGTCCTGGGTGTAGACCAGAACCACAGGTAAAAAAGAATGACTTCATT
AAATTAGCAGACAAATGGGTGGAACTAGAAAATGTCATCCTGGGCTGGAGA
GATGGCTCAGTGGTTCAGACCACTGGCTGCTCTTCCAGAGGTCCTGAGTTC
AATTCCCAACAACTATATGGTGGCTACCAACCATTACAATGAGATCAGATG
CCCTCCTCTTGTGTATCTGAAGAGAGTGACAGTGTACTTACATACATAAAA
TAAATAAATAAATCTAAAAAAAATGTTAAAAAA
>rGR05 aa
MLGAMEGVLLSVATSEALLGIVGNT
FIALVNCMDCTRNKNLYNIGFILTG
LAISRICLVWILITEAYIKIFSPQL
LSPINIIELISYLWIITSQLNVWFA
TSLSIFYFLKIANFSHHIFLWLKRR
INIVFAFLIGCLLMSWLFSFPVVVK
MVKDKKMLYINSSWQIHMKKSELII
NYVFTNGGVFLLFIIMLIVCFLLII
SLWRHSKWMQSNESGFRDLNTEVHV
KTIKVLLSFIILFILHLIGITINVI
CLLVPENNLLFVFGLTIAFLYPCCH
>rGR05 nt
AAGAGATTTCAGATACTACCACAAACATTTTTTAAATATATGTAAGTCTTT
AAAGAAAGAAGGGAAAGCCACTCCTTTATTGAGCAGCCAATAGATTGCCAT
CTTAAAATTCTGTGGCAGAAGCTATTTTAAAGATCTGCGAAGATGCTGGGT
GCAATGGAAGGTGTCCTCCTTTCAGTTGCAACTAGTGAGGCTTTGCTTGGC
ATTGTAGGGAACACATTCATTGCACTTGTGAACTGCATGGACTGTACCAGG
AACAAGAATCTCTATAATATTGGCTTCATTCTCACTGGCTTGGCAATTTCC
AGAATCTGCCTCGTGTGGATCTTAATCACAGAGGCATACATAAAAATATTC
TCTCCACAGTTGCTGTCTCCTATCAACATAATTGAACTCATCAGTTATCTA
TGGATAATTACCAGTCAATTGAATGTTTGGTTTGCTACCAGCCTCAGTATC
TTTTATTTCCTCAAGATAGCAAATTTTTCCCACCACATATTTCTCTGGTTA
AAAAGAAGAATTAATATAGTTTTTGCCTTCCTGATAGGGTGCTTACTTATG
TCATGGCTATTTTCTTTCCCAGTAGTTGTGAAGATGGTTAAAGATAAAAAA
ATGCTGTATATAAACTCATCTTGGCAAATCCACATGAAGAAAAGTGAGTTA

SLILILANSRLKRCFV[redacted]QQLMCS EEGKEFRNT	ATCATTAACTATGTTTTCACCAATGGGGG[redacted]ATTTTTACTTTTTATAATA ATGTTAATTGTATGTTTTCTCTTAATTAT[redacted]CCTTTGGAGACACAGCAAG TGGATGCAATCAAATGAATCAGGATTCAGAGATCTCAACACAGAAGTTCAT GTGAAAACAATAAAAGTTTTATTATCTTTTATTATCCTTTTTATATTGCAT TTAATTGGTATTACCATCAATGTCATTTGTCTGTTAGTCCCAGAAAATAAC TTGTTATTCGTGTTTGGTTTGACGATTGCATTCCTCTATCCCTGCTGCCAC TCACTTATCCTAATTCTAGCAAACAGCCGGCTGAAACGATGCTTTGTAAGG ATACTGCAACAATTAATGTGCTCTGAGGAAGGAAAAGAATTCAGAAACACA TGACAGTCTGGAAGACAAACAATCAGAAATAGTAAGTGAAAAAAAAAAAAA AAAAA
>rGR06 aa (partial) EALVGILGNAFIALVNFMGWMKNRK ITAIDLILSSLAMSRICLQCIILLD CIILVQYPDTYNRGKEMRIIDFFWT LTNHLSVWFATCLSIFYFFKIANFF HPLFLWIKWRIDKLILRTLLACLIL SLCFSLPVTENLADDFRRCVKTKER INSTLRCKLNKAGYASVKVNLNLVM LFPFSVSLVSFLLLILSLWRHTRQM QLNVTGYNDPSTTAHVKATKAVISF LVLFIVYCLAFLIATSSYFMPESEL AVIWGELIALIYPSSHSFILILGNS KLKQASVRVLCRVKTMLKGRKY	>rGR06 nt (5'-truncated) GTGAGGCCTTAGTAGGAATCTTAGGAAATGCATTCATTGCATTGGTAAACT TCATGGGCTGGATGAAGAATAGGAAGATCACTGCTATTGATTTAATCCTCT CAAGTCTGGCTATGTCCAGGATTTGTCTACAGTGTATAATTCTATTAGATT GTATTATATTGGTGCAGTATCCAGACACTTACAACAGGGGTAAAGAAATGA GGATCATTGATTTCTTCTGGACGCTTACCAACCATTTAAGTGTCTGGTTTG CCACCTGCCTCAGCATTTTCTATTTCTTCAAGATAGCAAACTTCTTCCATC CTCTTTTCCTCTGGATAAAGTGGAGAATTGACAAGCTAATTCTGAGGACTC TACTGGCATGCTTGATTCTCTCCCTATGCTTTAGCCTCCCAGTCACTGAGA ATTTGGCTGATGATTTCAGACGCTGTGTCAAGACAAAAGAAAGAATAAACT CTACTCTGAGGTGCAAATTAAATAAAGCTGGATATGCTTCTGTCAAGGTAA ATCTCAACTTGGTCATGCTGTTCCCCTTTTCTGTGTCCCTTGTCTCATTCC TTCTCTTGATTCTCTCCCTATGGAGACACACCAGGCAGATGCAACTCAATG TAACAGGGTACAATGATCCCAGCACAACAGCTCATGTGAAAGCCACAAAAG CAGTAATTTCCTTCCTAGTTCTGTTTATTGTCTACTGCCTGGCCTTTCTTA TAGCCACTTCCAGCTACTTTATGCCAGAGAGTGAATTAGCTGTAATTTGGG GTGAGCTGATAGCTCTAATATATCCCTCAAGCCATTCATTTATCCTGATCC TTGGGAACAGTAAACTAAAACAGGCATCTGTAAGGGTGCTTTGTAGAGTAA AGACTATGTTAAAGGGAAGAAAATATTAGCATCATGGATATATTTGAAGAA AAACTATCACTGTCTAAAGAAAAAGGATGACAAATCATTATCTTTCATTCT TATATGAATATTGCTTTCATGCGGTAACATCTTTTAACAAACTTAAATCAA ATGTTGGGAAATCTCATATACAGCAACTTTGCATGTCTCTCTGTCTATTTC CCTCTCCCTTTGTACATAGTTGACATAAAAAAAGAATTTTCATGACAAAAT TGTAATAAATAGCTACAGAGGCAGCACATTTTCATAGTAAGTTCTGAATCA CTCTTCCAAATGCAAAGCTGCCTGACAAATTCAAAACAACTGTAACAGTAT TTCACTGCTGTTTGCATTCTTTGGAAAAGCAGGTGGTTTGTTCCTATGACC TGACTTGGAGTTTTCTTCTTACATCACTG
>rGR07 aa MGSSLYDILTIVMIAEFIFGNVTNG FIVLTNCIAWLSKRTLSFIGWIQLF LAISRVVLIWEMLLAWLKYMKYSFS YLAGTELRVMMLTWVVSNHFSLWLA TILSIFYLLKIASFSRPVFLYLKWR VKKVLLLILLGNLIFLMFNILQINT HIEDWMDQYKRNITWDSRVNEFVGF SNLVLLEMIMFSVTPFTVALVSFIL LIFSLWKHLQKMHLSSRGERDPSTK AHVNALRIMVSFLLLYATYFISFFI SLIPMAHKKGLDLMFSLTVGLFYPS SHSFILILGHSNLRHSSCLVITYLR CKEKD	>rGR07 nt CAGTAGCAAAATTTTACTATGTTCATTGATATTATGTCAnGnCACTACGTA AGAAGGAAGACTTGAAAGAAAGCTTATCTGAGTTTTTAAGAATACATGGAC ATTTCAGCTTGGCAAATGACGAGCTGTGAATTTTTGTCATCTGGACATGGG AAGCAGCCTGTATGATATCTTAACTATTGTCATGATTGCAGAGTTTATATT CGGAAATGTGACCAATGGATTCATAGTGCTGACAAACTGTATTGCTTGGCT CAGTAAAAGAACTCTTTCTTTCATTGGTTGGATCCAGCTTTTCTTGGCCAT TTCCAGAGTGGTTTTGATATGGGAAATGTTACTAGCATGGCTGAAATATAT GAAGTATTCATTTTCATATTTGGCTGGCACAGAATTAAGGGTTATGATGTT GACCTGGGTAGTTTCCAATCACTTTAGTCTCTGGCTTGCCACCATTCTAAG CATCTTTTATTTGCTCAAAATAGCTAGTTTCTCCAGACCTGTTTTCCTGTA TCTGAAGTGGAGAGTAAAAAAAAGTGCTCCTGCTGATTCTTCTCGGAAATTT AATCTTCCTGATGTTCAATATATTACAAATCAACACTCACATAGAAGACTG GATGGATCAATATAAGAGAAATATAACGTGGGATTCCAGAGTGAATGAATT TGTGGGGTTTTCAAATCTGGTTTTATTGGAGATGATTATGTTCTCTGTAAC ACCATTCACCGTGGCTCTGGTCTCCTTCATCCTGTTAATCTTCTCTTTATG GAAACATCTCCAGAAGATGCATCTCAGTTCCAGAGGGGAACGAGACCCTAG CACAAAAGCCCATGTGAATGCCCTGAGAATTATGGTCTCCTTCCTCTTACT CTATGCCACTTACTTCATATCCTTTTTTATATCATTAATTCCTATGGCACA TAAAAAAGGACTAGATCTTATGTTTAGCCTAACTGTTGGACTTTTCTACCC TTCAAGCCACTCATTTATCTTGATTTTGGGACATTCTAATCTAAGGCATTC CAGTTGTCTGGTGATAACCTATCTGAGATGTAAGGAAAAGGATTAGAAATT CACTATTCCATAAGGCAGTTAAACCACATGCTATTAGGTATACTCAGTGCT AGATCCCTAGGCAAGCATTAACATTAAAAATATATAATTTCTAGATTCTTC TATTTGTGATAAACCACTCACTTAGAATAATGCTAAAGTAGCGTGATGTTG TATATAAGTGTAAGAATAAAATGTAATTAATTTAGTTTAGGCACAATAACA



	TATGTCTACTAAGTAAAAACTAGGCAGG[redacted]CTACACGCATATTAGAATCC AGGCTGAGGTATATAGACTCAAGAAATAC[redacted]GTGGAATAAAGATTTTAATTT TCATTCTATTGTGAGTTATGTGAAATCAATGCCATTAAAGGCATACACAAG ATTTTCACACACTGAAACAACTTCTTGCATTTTGTCATATTGTATTGGAAG TAAATTGGAGATAAACTTAATATCAATAAATTACAAAATGTAAACATAAAC AGGGTGATTAAAAATTAGCCTCTAGGTCCTGGGGAAATGATTCaAGTAAAG TGCTTTCTTTTCAAATAGGAGAATCTGATTGTAAATCATCTAAAAGTCTGG CATAAAATGTCAATGAAAATTGTATGTAAAATATAGCTATgGCmAAGAGCA CCmAAGAAAAGAAAATTTTTGCCTTTGAAACCCAGTAATTGATATCCTTTA AAAAAGCAGTTACATATTTTTCTGTTTAAGATTTTGTCAAAGGGTAGCTTT GACAACTAATATAAGCTGAGGAAGGTAGCAAGTGTGAAGTCAGCTAATGGG GTCAGTCAAGTGCTGTTAGCAGCAGATGGAGGCCACTGCTGAATTTAGCAG GCAATTTACAGGGTGAGCACTGCTAGTGCTGACAGAAGAAAAACTCTGAAA TTTTAACTCTTTAGGGTCTGGTGAGAAAGAAAAAGAGAGAAAATCGCATAT ATATATATATATATATATATATATATATATATATATATATATATATATATA TCATGGAAGCTCTAACAAGTTGACTCAAACAACTTTATGATGTTTTTAGGC CCTTTTATTTTAATGTCAGTGAATTAGGTGTGGTACAGCAATATTGCTACT TTTAAATTCAAAGCAGTTGTTTTATATATTATTCATTATATAAGCTAATTA TAAGTTTAAATCAAAAGGTTTATTTGTCCATGATTTTACTTTATCATTGGG CACACCTGTGCTCTCATCCTTGGGCTTGACCTAGAATGAAAGTTTATCCTT GATCATATGTCTGTCACAAGACTACTTCTCTTCCTATAGTAGTTTATGTAC TTACAATATACAAAAGTTTATTGAATTCCTTTTATCACTTATGCAGCCTTT TCTTACTATTCTATTCTATTCTATTCTATTCTATTCTATTCTATTCTATTC TATTCTATTCTATTCTATTCTATTCTATTCTAGAATCTAACCTATACATTC ATTTCTGGCAAAACAACTTATATCATCTCCTTAATTATTTTATCAATTAAT CTAACATCCTGAAGTTATTTAAATCTAATATAAGGACTCTGTAAAGTCACA AATTTATTTATACTTCACAAAATTCATTATTTTATGGAACTGCAGCATTGC CTGGGCCAGGAGTCACAAGAGTTCCAGAGTTGACTTTATTGGCATCTGCCT GGCTAACTGAAGGATCAGTTTTCTGTGTACAATAATTTTGTGTATCTCTTT TGATGCAAGATATGAAAAATAATTTCAGTCTAAAAGTGTCCTTAAATTTGA AACTCTCTGGCCAGAATCTAACTATTGATGACCAGTTTGCACCATGGACTC AGTGTCTTCTATTGCTTTAAAATAAGCAACATCTTGAATGCTTTTCTTGTG TATTAGGCAAATAATTAACAACATGTTTCTATGATTGTCTCAATAACAATA CTATATTTCTCACAGTTTTTAATTTTTATGGCAAAGTTGGCTAATAAGAAT TTTTTTCAAATTATCAAACGTGAAGAAAACTTGACATTTTATTTCATGGAG ATTCTAAATGTTTTCTTAGCATATTGCCTTTTTACTAACTTGATTTTTATC ATGTTTTGGTAGTATTTCTAATTTTCCTTTTTTTCTAAGTATGTTATGTAG TAACACCAGGAGAATGAAACAAATGACATTTATACTAAGGATGTGACAAAT AAGGCCCAAAGAAAGTTTTGAAAATCATGATCTCATTTCTATTCTTCTTTA TTAAGTATAGCATAAGCAAAATTCTGATGGTGGTCTTGGCCCATATCTTTG AACACAGTGTAGTGGTGAAGACTTTTTCAAATATTATGTCATATTTGTACC CATCTCTGTACCTATTTCTTCTGATTTCATGAGGAAAAAATGAGGAAGGGT TTGTTTGTGTGCTGGAGCAGCTGAAGTGGACCAAGGGGCAGGAATTCTCTC TGTTCGGTCCTAGTGTGACTGATGATGCTCTCATTGAAAAACAGGAAGAAG AAGAAAGACTTTATATGCACCATTCACTCCTTCCCCCTCCTACATTCCACC TCCCTCTTGAAAGAGTGTCTATCTATATAGATATAGCTATCCTGAAATCCA TTAAGTAGACCTGACTGGCTTAAATCTCACAGAAATTCACCTACCTTTTCC ATGATTGCTGAAATTAAAGACATGTGCCGACATATTGGGCACATTCAGACC TTTTGCCAACTGTCTTTCAACTCATTTGGACCTACTGAGAAGTATTCAAAA TATTTGGTTGTTTTAAATAAAAGGAAAGTGGGTCTATATTACTTGAATTGG ATAGAGAAATTTTCACTTACAAGTGATATTGAAAATGGGGGAGAATGTATT TTAGCATAAGCACCAGAACACAAAGCAATTCTTGTTAAAACTTTATCGATA AATTGGATAAATGTTAAAAAAGAAAAAATAAAATATACGAACTATTATGAA AAAAAAAAAAAAAAAA
>rGR08 aa MEPVIHVFATLLIHVEFIFGNLSNG LIVLSNFWDWVVKRKLSTIDKILLT LAISRITLIWEMYACFKIVYGSSSF IFGMKLQILYFAWILSSHFSLWFAT ALSIFYLLRIANCSWKIFLYLKWRL KQVIVGMLLASLVFLPGILMQRTLE ERPYQYGGNTSEDSMETDFAKFTEL	>rGR08 nt CTGCAGGTTGGTGATCCAGTAATGAGCAGCACTGTTATATCTCAGGCTTTC TAAGATCATGGAACCTGTCATTCACGTCTTTGCCACTCTACTAATACATGT GGAGTTCATTTTTGGGAATCTGAGCAATGGATTAATAGTGTTGTCAAACTT CTGGGACTGGGTCGTTAAACGAAAACTTTCCACAATTGATAAAATTCTTCT TACATTGGCAATTTCAAGAATCACTCTCATCTGGGAAATGTATGCTTGTTT TAAAATTGTATATGGTTCATCTTCATTTATATTTGGGATGAAGTTACAAAT TCTTTATTTTGCCTGGATCCTTTCTAGTCACTTCAGCCTCTGGTTTGCCAC AGCTCTCAGCATCTTTTACTTACTCAGAATAGCTAACTGCTCCTGGAAGAT

ILFNMTIFSVIPFSL[redacted]SFLLLIF SLWKHLQKMQLSSRGHGDPSTKAHR NALRIMVSFLLLYTSYFLSLLISWI AQKHHSKLVDIIGIITELMYPSVHS FILILGNSKLKQTSLWILSHLKCRL KGENILTPSGKPIN	CTTCCTGTATCTGAAATGGAGACTTAAA[redacted]GTGATTGTGGGGATGTTGCT GGCAAGCTTGGTGTTCTTGCCTGGAATCC[redacted]ATGCAAAGGACTCTTGAAGA GAGGCCCTATCAATATGGAGGAAACACAAGTGAGGATTCCATGGAAACTGA CTTTGCAAAGTTTACAGAGCTGATTCTTTTCAACATGACTATATTCTCTGT AATACCATTTTCATTGGCCTTGATTTCTTTTCTCCTGCTAATCTTCTCTTT GTGGAAACATCTCCAGAAGATGCAGCTCAGTTCCAGAGGACATGGAGACCC TAGCACCAAGGCCCACAGAAATGCTTTGAGAATTATGGTCTCCTTCCTCTT GCTCTACACTTCATATTTCCTGTCTCTTCTTATATCATGGATTGCTCAGAA GCATCACAGTAAACTGGTTGACATTATTGGTATTATTACTGAACTCATGTA TCCTTCAGTCCACTCATTTATCCTGATTCTAGGAAATTCTAAATTAAAGCA GACTTCTCTTTGGATACTGAGTCATTTGAAATGTAGACTGAAAGGAGAGAA TATTTTAACTCCATCTGGCAAACCAATTAACTAGCTGTTATATATTCTGTA TTGCAAACAAATCAGTGAGTTAGTGGTTCAAGGATTCCATCCTTGACTTAT TGTATCATGGAAGTCATATAGGGAGAGGCTGAACAAGCTATCTTCTGTAAA TTGGCAAGGGTTGCATATAGTACTGGTACTGGGACACCATCCAACCATAAA ACCTTCTAACCATAACCTACCTGACTGCAAGATATGCTGGGACAATGGTGG CTCAGAGATTTTGGGACTGGCCAACCAATGTCTATTCTTTCTTGAGGCTCA CTCAATAAGGAGGCCATGCCCAACTCGTCcTGGATGGCCAGGAACCAGAAT CTCTGATGGsCCAATGATCTATGGnAGAACCCAGCATTACTGGGAAAAAAG AATAATCACTTTGATGAATGGTCAAATATTTCCTAAATATATTCTGATACA CTTGTACATCATTTCTCTTTCCCAATCATCATCACAGGGACTTCTCCCCAG CACCTGATGGGAACAGATACCAAAATCTACAGCCAAATACTAAATGCAGGT TGGGGAACTCCACAAAAGACTGGAAGGAAGTACTGTGAGAGCCAGAGTGGT CCAGAACACTAGGAGAACACAGAACATCGAATTAACTAAGCAGCACTCATA GGGTTAATGTAAAATAAAGCAGCAGTCACATAGACTGCACAGGTGTACTCT AGATCCTCTGCATATATGTTGTGGTTGTCAAACTTGGGAGTTTTGTTGGAC TAATAACAATGTGAATAAGTAAGTCTCTGACACTTATTCCCGCTCTTGGAA CCCTTTTCCACATTTTGTATTGTCTTACCACCTTGATATGAAGGTTTCTGA ATAGTCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
>rGR09 aa MLSAAEGILLSIATVEAGLGVLGNT FIALVNCMDWAKNKKLSKIGFLLFG LATSRIFIVWILILDAYAKLFFPGK YLSKSLTEIISCIWMTVNHMTVWFA TSLSIFYFLKIANFSHYIFLWLKRR TDKVFAFLLWCLLISWAISFSFTVK VMKSNPKNHGNRTSGTHWEKREFTS NYVLINIGVISLLIMTLTACFLLII SLWKHSRQMQSNVSGFRDLNTEAHV KAIKFLISFIILFILYFIGVAVEII CMFIPENKLLFIFGLTTASVYPCCH SVILILTNSQLKQAFVKVLEGLKFS ENGKDLRAT	>rGR09 nt GGACACTGCAGCAGATCTGCTATAGAATAACAGATACAAACATAGCAACCT GCAGAGATGCTCAGTGCAGCAGAAGGCATCCTTCTTTCCATTGCAACTGTT GAAGCTGGGCTGGGAGTTTTAGGGAACACATTTATCGCCCTGGTTAACTGC ATGGATTGGGCCAAGAACAAGAAGCTCTCTAAGATTGGTTTCCTTCTCTTT GGCTTAGCAACTTCCAGAATTTTTATTGTATGGATATTAATTTTAGACGCA TATGCAAAGCTATTCTTTCCGGGGAAGTATTTGTCTAAGAGTCTGACTGAA ATCATCTCTTGTATATGGATGACTGTGAATCACATGACTGTCTGGTTTGCC ACCAGCCTCAGCATCTTCTATTTCCTAAAAATAGCAAATTTTTCCCACTAT ATATTTCTCTGGTTAAAGAGGAGAACTGATAAAGTATTTGCCTTTCTCTTG TGGTGTTTATTAATTTCATGGGCAATCTCCTTCTCATTCACTGTGAAAGTG ATGAAGAGCAATCCAAAGAATCATGGAAACAGGACCAGTGGGACACATTGG GAGAAGAGAGAATTCACAAGTAACTATGTTTTAATCAATATTGGAGTCATT TCTCTCTTGATCATGACCTTAACTGCATGTTTCTTGTTAATTATTTCACTT TGGAAACACAGCAGGCAGATGCAGTCTAATGTTTCAGGATTCAGAGATCTC AACACTGAAGCTCATGTGAAAGCCATAAAATTTTTAATTTCATTTATCATC CTTTTCATCTTGTACTTTATAGGTGTTGCAGTAGAAATCATCTGCATGTTT ATCCCAGAAAACAAACTGCTATTTATTTTTGGTTTGACAACTGCATCCGTC TATCCCTGCTGTCACTCAGTCATTCTAATTCTAACAAACAGCCAGCTGAAG CAAGCCTTTGTAAAGGTACTGGAGGGATTAAAGTTCTCTGAGAACGGAAAA GATCTCAGGGCCACATGAGTCTGGAACAGAAATGGGTAGTCTGGAATAATT GTAAGGAAGTCGTAGAAGGTCTTTTTCATTTGTACAGTGCTCTTACCTTGT TTTTGAGGAGATGTAAACTTTTTTATTTTTATTTTTTATCCTATGTGAATA AGTGTGTGTGTGTGTGTGTGTTTATGTGTGTGTGTATATATGTCTATGT GTGTTTTAGGAGGTTTAAGAGGGAAGAGGGAATAGAGGTATGTTGGTGTTT TTAACATGGATATTCACAGGCCAAGGAACTTGTTCTCTCCTTTTACCTTAG GGTAGTGTCCTTTGTGGCTGTCACTCTGACAGTCTACACTAGTTGAACTAA GAGCTTTTAGCCAGTTCACTTGTCTAAACCTCCCTTCTCATGGTAGCAGTG TTCTGATTACAGAATCATGCTGTCACATACAGCTTTTTTAACAAGGTTCCCA TAGACAGAATTCATGTCAAACGGAATGCACAGCTGTCACTCTTACCCACCG ATCTCTCTTGCCAGCCCATTCCTATTGACTTTAAACTGTAGTATTAAACTT TACTGAAATCTTCTGCAACCAGTCTGACTATGTCTCTTGAAATCACATGAT ATGGTGGAATTTTAATGCCATGTGAAAATTTGTTTGTTCAGTTAGTTTCCT ACTCTGCCAAATCATTCTCTTACACTTGGCAGAAAAAAACCATCAACTGTA

	GACTATTTTGTGTAAAGACTAATACAGAT[redacted]ATAAGTATCTTAATCAAGA TGTCATTGTGATTATCCTAATTTCCCCAG[redacted]ACTGGTTCCCTTTCCCCAG AAAGACTCACAAAGGAACTGAGGCAAACAGTTGTGGTCACTCTTGATATTT ACCAGTTGAAACTGAAGAACAGTGTTTCCTTTCTGTTCAGTTTTACTACTT ACAGTTACTTTATTTCATCCATTAAATCCCAAAGTGCTTATTAATAGTAGA TATTTGATGAAGCAACAATGGTTATAAGAGTGGATGTGGATCTATGACAAA GATCTAGAGAAACAGACTATTTGTGAAAGATGGATGAAAGCCCTGATGAAA GGATTCTTCATGGTCTTTGACCCCAGGGAGTTTTGAAATCAAGCAGCCACA GATCAAAGAGAGCTGAGAAGAGGTTCTCCTGAAGAAAATATCCAAACACAT GGTGCCAGCCAAAGCAGAAAATAGTGGACAATTCAGTCCAGGACCTGAATG AGGTAGACAATGTCCTGTTAAGGGTTGGAACAAATATATAGATATGGTCAT TCATATACAGAAACCTACAGGCGTGTTTGAACTCTTGGTTTCTCAGTAATC AATTCTTAAATCTTTTTTAGAATGGATTTTTTATCATCATTCATGATCTCT CAGCAGAGTCTGCAGGGGCTAAGAGACACACTAAGAGTATCTGGAGGGGGG AGTGTCTTCCTGCTCTATCAACCCCTAAAGTCATATATAACAATACAAAAT TCCACATTAGTTAAGTTCTTTTTTTTACATCTTTATTAAATTGGGTATTTC TTATTTACATTTCAAATGTGATTCCCTTTCCTGGTTTCCAGGCCAATATCC CCCTAACCTCTCCCCTTCTATGTGGGTATTCCCTCGTGCCGAATTC
>rGR10 aa (partial) MFLHTIKQRDIFTLIIIFFVEITMG ILGNGFIALVNIVDWIKRRRISSVD KILTTLALTRLIYAWSMLIFILLFI LGPHLIMRSEILTSMGVIWVVNNHF SIWLATCLGVFYFLKIANFSNSLFL YLKWRVKKVVLM	>rGR10 nt (3'-truncated?) CCCGGGCTGCAGGATTCGGCACGAGAATGAAAACTTTTGCTCTACTATTTT GCTGTTCTGTGATACCACAGACCATAAAACAATCGAGCCAAGGGATCAAGA GCTGAAACTTCAGAAAGTGGGAATCAAATTTCCTTCCTGATAGGTTAGCTT ATGAGAATTCAGCATCTTATTCAACTTCAGAAAATTGGATATAAGATACAG TGTCTGGATGAAGCCGAATTGATCTATTTGGGGAGAAAAAACGCCAACATT TATAATAAGGTTTTATGAGACAGTTCCTGGGAAATTTGGATATTTCCTAGT TAGTAATGTGTAAATGGGATTTTAAAACATGATTATTTTGTATTTTTAACA ACCAACATGAGGAGCTTTTTAAATGCCACTTAGACATTATAAACTGAAGCA TGTTCTTACACACAATAAAGCAACGTGATATTTTTACTTTGATAATCATAT TTTTTGTGGAAATAACAATGGGAATCTTAGGAAATGGATTCATAGCACTAG TGAACATTGTGGACTGGATCAAGAGAAGAAGGATTTCTTCAGTGGATAAGA TTCTCACTACCTTGGCCCTTACCAGACTCATTTATGCGTGGTCTATGCTCA TTTTTATATTGTTATTCATACTGGGCCCGCATTTGATTATGAGATCAGAAA TACTTACATCAATGGGTGTTATCTGGGTGGTGAACAATCACTTCAGCATCT GGCTTGCTACATGCCTCGGTGTCTTTTATTTTCTCAAGATAGCCAATTTTT CTAACTCTTTGTTTCTTTACCTAAAGTGGAGAGTTAAAAAAGTGGTTTTAA TG ... poly(dA)???
>rGR11 aa GSGNGFIVSVNGSHWFKSKKISLSD FIITSLALFRIFLLWIIFTDSLIIV FSYHAHDSGIRMQLIDVFWTFTTHF SIWLISCLSVFYCLKIATFSHPSFL *LKSR	>rGR11 nt GGATCCGGAAACGGTTTTATCGTGTCAGTCAATGGCAGCCATTGGTTCAAG AGCAAGAAGATTTCTTTGTCTGACTTCATCATTACCAGCTTGGCCCTCTTC AGGATCTTTCTGCTGTGGATCATCTTTACTGATAGCCTCATAATAGTGTTC TCTTACCACGCCCACGACTCAGGGATAAGGATGCAACTTATTGATGTTTTC TGGACATTTACAACCCACTTCAGTATTTGGCTTATCTCCTGTCTCAGTGTT TTCTACTGCCTGAAAATAGCCACTTTCTCCCACCCCTCATTCCTGTAGCTC AAATCTAGA
>rGR12 aa MLSTVSVFFMSIFVLLCFLGILANG FIVLMLSREWLWRGRLLPSDMILLS LGTSRFCQQCVGLVNSFYYSLHLVE YSRSLARQLISLHMDFLNSATFWFG TWLSVLFCIKIANFSHPAFLWLKWR FPALVPWLLLGSILVSFIVTLMFFW GNHTVYQAFLRRKFSGNTTFKEWNR RLEIDYFMPLKLVTTSIPCSLFLVS ILLLINSLRRHSQRMQHNAHSLQDP NTQAHSRALKSLISFLVLYALSYVS MVIDATVVISSDNVWYWPWQIILYL CMSVHPFILITNNLKFRGTFRQLLL LARGFWVT	>rGR12 nt GTGTGAGGGACTGTGGGTAGGGGCTGGGAGGAGGCCAGGAACCAAGGCAAC CAGTGGTGACAGGAGGGGCTGAAATGCTATCAACTGTATCAGTTTTCTTCA TGTCGATCTTTGTTCTGCTCTGTTTCCTGGGAATCCTGGCAAACGGCTTCA TTGTGCTGATGCTGAGCAGGGAATGGCTATGGCGCGGTAGGCTGCTCCCCT CAGACATGATCCTCCTCAGTTTGGGCACCTCCCGATTCTGCCAGCAGTGCG TTGGGCTGGTGAACAGTTTCTACTATTCCCTCCACCTTGTTGAGTACTCCA GGAGCCTTGCCCGTCAACTCATTAGTCTTCACATGGACTTCTTGAACTCAG CCACTTTCTGGTTTGGCACCTGGCTCAGCGTCCTGTTCTGTATCAAGATTG CTAACTTCTCCCATCCTGCCTTCCTGTGGTTGAAGTGGAGATTCCCAGCAT TGGTGCCTTGGCTCCTACTGGGCTCTATCTTGGTGTCCTTCATCGTAACTC TGATGTTCTTTTGGGGAAACCACACTGTCTATCAGGCATTCTTAAGGAGAA AGTTTTCTGGGAACACAACCTTTAAGGAGTGGAACAGAAGGCTGGAAATAG ACTATTTCATGCCTCTGAAACTTGTCACCACGTCAATTCCTTGCTCTCTTT TTCTAGTCTCAATTTTGCTGTTGATCAATTCTCTCAGAAGGCATTCACAAA GAATGCAGCACAATGCTCACAGCTTGCAAGACCCCAACACCCAGGCTCACA GCAGAGCCCTGAAGTCACTCATCTCATTTCTGGTTCTTTACGCGCTGTCCT

	ATGTGTCCATGGTCATTGACGCTACAGT[redacted]CATCTCCTCAGATAACGTGT GGTATTGGCCCTGGCAAATTATACTTTAC[redacted]GTGCATGTCCGTACATCCAT TTATCCTTATCACTAATAATCTCAAGTTCCGAGGCACCTTCAGGCAGCTAC TCCTGTTGGCCAGGGGATTCTGGGTGACCTAGAAGGTTTGGTCTCTTTATC TGTACCCTTTGAAGAGACTTAGGTGAGGGTGACTTCCCTTGGAAGTGATCT CATCTACATGGAAATGTCTTTGTAGGCTGACATGGGGTCATACTATGTGGT TCCTCCTTGGGAAAGAGGAGAAGAAAATACAGGGATTCTGAGCGTTCTTCC TTATCTTGGGATATTATGAAAATGGACATTCTGAATCCTGAACCAGTATTG ATCTGAAGTGCAAAGTACAATATGCCTGTTCCCTTCATGTCTGCTATCCTC TTGGTACTTATTAATTCCCT … approximately 500 bp to end
>rGR13 aa MCGFPLSIQLLTGLVQMYVILIIAV FTPGMLGNVFIGLVNYSDWVKNKKI TFINFILICLAASRISSVLVVFIDA IILELTPHVYHSYSRVKCSDIFWVI TDQLSTWLATCLSIFYLLKIAHFSH PLFLWLKWRLRGVLVGFLLFSLFSL IVYFLLLELLSIWGDIYVIPKSNLT LYSETIKTLAFQKIIVFDMLYLVPF LVSLASLLLLFLSLVKHSQNLDRIS TTSEDSRAKIHKKAMKMLLSFLVLF IIHIFCMQLSRWLFFLFPNNRSTNF LLLTLNIFPLSHTFIIILGNSKLRQ RAMRVLQHLKSQLQELILSLHRLSR VFTMEIA	>rGR13 nt GGGATTCAGTTGGATAAGAGAAAAGTCAAAACCCTAAGACTAAGAATTTCC TTAAGTAGATATCAATTTCTATCCATTGGAAGGAGTTTCCAATCACACTGA AATTACAATAAAAAAGGAGCAAGATAACTATGGGAAAGGATGATTTTCGGT GGATGTTTGAGAACTGAGCAGCAAGGCAAATTGATAGATGTGTGGATTCCC TCTTTCTATTCAACTGCTTACTGGATTGGTTCAAATGTACGTGATATTGAT AATAGCAGTGTTTACACCTGGAATGCTGGGGAATGTGTTCATTGGACTGGT AAACTACTCTGACTGGGTAAAAAACAAGAAAATCACCTTCATCAACTTCAT CCTGATCTGTTTGGCAGCGTCCAGAATCAGCTCTGTGTTGGTGGTATTTAT TGATGCAATCATCCTAGAACTAACTCCTCATGTCTATCATTCTTACAGTCG AGTGAAATGCTCTGATATATTCTGGGTTATAACTGACCAGCTGTCAACGTG GCTTGCCACCTGCCTCAGCATTTTCTACTTACTCAAAATAGCCCACTTCTC CCATCCCCTTTTCCTTTGGTTGAAGTGGAGATTGAGAGGAGTGCTTGTTGG TTTTCTTCTATTTTCTTTGTTCTCATTGATTGTTTATTTTCTACTCCTGGA ATTACTGTCTATTTGGGGAGATATTTATGTGATCCCTAAAAGCAATCTGAC TTTATATTCAGAAACAATTAAGACCCTTGCTTTTCAAAAGATAATTGTTTT TGATATGCTATATTTAGTCCCATTTCTTGTGTCCCTAGCCTCATTGCTCCT TTTATTTTTATCCTTGGTGAAGCACTCCCAAAACCTTGACAGGATTTCTAC CACCTCTGAAGATTCCAGAGCCAAGATCCACAAGAAGGCCATGAAAATGCT ATTATCTTTCCTCGTTCTCTTTATAATTCACATTTTTTGCATGCAGTTGTC ACGGTGGTTATTCTTTTTGTTTCCAAACAACAGGTCAACTAATTTTCTTTT GTTAACATTAAACATCTTCCCATTATCTCATACATTCATTATCATCCTGGG AAACAGCAAGCTTCGACAAAGAGCAATGAGGGTCCTGCAACATCTTAAAAG CCAACTTCAAGAGTTGATCCTCTCCCTTCATAGATTGTCCAGAGTCTTCAC TATGGAAATAGCTTAAAGGGGAGACTTGGAAGGTCACTGGTAACTTGTTCT TCCGCTGAGTTCTGTTAAGTAATGCTGGACATATATGAACTATCCCTAGTG CATACTGATATT … approximately 1500 bp to end?
>rGR14 aa (partial) VANIMDWVKRRKLSAVDQLLTVLAI SRITLLWSLYILKSTFSMVPNFEVA IPSTRLTNLVWIISNHFN	>rGR14 nt (oligo sequence removed) CTGTGGCAAACATAATGGATTGGGTCAAGAGAAGGAAGCTCTCTGCAGTGG ATCAGCTCCTCACTGTGCTGGCCATCTCCAGAATCACTCTGTTGTGGTCAT TGTACATACTGAAATCAACATTTTCAATGGTGCCAAACTTTGAGGTAGCTA TACCGTCAACAAGACTAACTAATCTTGTCTGGATAATTTCTAACCATTTTA AT

>mGR01 aa (notional) MQHLLKTIFVICHSTLAIILIFELI IGILGNGFMALVHCMDWVKRKKMSL VNKILTALAISRIFHLSLLLISLVI FFSYSDIPMTSRMTQVSNNVWIIVN HFSIWLSTCLSVLYFLKISNFSNSF FLYLKWRVEKVVSVTLLVSLLLLIL NILLINLEISICIKECQRNISCSFS SHYYAKCHRQVIRLHIIFLSVPVVL	>mGR01 nt AGCTGTGCGTGAGCAAAGCATTTCTTGTCTGCCACTTCTGAGCTGTGTGAG GAGACACATTATCACGGAAAGAGATTCAGACTCTGTCGCTGTCAAACCTGT ATGTTTGCTCCTCTTTTACTGTGAAGGCAGAGTTACGAAAAAAAATGTTAT GAGAACCAACTCAGAAATTGACAAAAATTTTCTAAATGTCATTTTTAAAAA TTATATTTCAAATGGAAATGTGAGCAAATCTTTATAACTAATATATAAAAT GCAGCATCTTTTAAAGACAATATTTGTTATCTGCCATAGCACACTTGCAAT CATTTTAATCTTTGAATTAATAATTGGAATTTTAGGAAATGGGTTCATGGC CCTGGTGCACTGTATGGACTGGGTTAAGAGAAAGAAAATGTCCTTAGTTAA TAAAATCCTCACTGCTTTGGCAATCTCCAGAATTTTTCATCTCAGTTTATT

SLSTFLLLIFSLWTL[redacted]MQQHVQG GRDARTTAHFKALQTVIAFFLLYSI FILSVLIQNELLKKNLFVVFCEVVY IAFPTFHSYILIVGDMKLRQACLPL CIIAAEIQTTLCRNFRSLKYFRLCC IF	GCTTATAAGTTTAGTCATATTCTTTTCA[redacted]TCTGATATTCCTATGACTTC AAGGATGACACAAGTCAGTAATAATGTT[redacted]GATTATAGTCAATCATTTCAG TATCTGGCTTTCTACATGCCTCAGTGTCCTTTATTTTCTCAAGATATCCAA TTTTTCTAACTCTTTTTTTCTTTATCTAAAGTGGAGAGTTGAAAAAGTAGT TTCAGTTACACTGTTGGTGTCATTGCTCCTCCTGATTTTAAATATTTTATT AATTAACTTGGAAATTAGCATATGCATAAAGGAATGTCAAAGAAACATATC ATGCAGCTTCAGTTCTCATTACTATGCAAAGTGTCACAGGCAGGTGATAAG GCTTCACATTATTTTCCTGTCTGTCCCCGTTGTTTTGTCCCTGTCAACTTT TCTCCTGCTCATCTTCTCCCTGTGGACACTTCACCAGAGGATGCAGCAGCA TGTTCAGGGAGGCAGAGATGCCAGAACCACGGCCCACTTCAAAGCCCTACA AACTGTGATTGCATTTTTCCTACTATATTCCATTTTTATTCTGTCTGTCTT AATACAAATATGAATTACTGAAGAAAAATCTTTTCGTTGTATTTTGTGAGG TTGTATATATAGCTTTTCCGACATTCCATTCATATATTCTGATTGTAGGAG ACATGAAGCTGAGACAGGCCTGCCTGCCTCTCTGTATTATCGCAGCTGAAA TTCAGACTACACTATGTAGAAATTTTAGATCACTAAAGTACTTTAGATTAT GTTGTATATTCTAGACAAAAATTAACTGATACAAATGTCTTTTGTATTTTT CATTTTAAATATCCTTTAATTTTGACTGCATGAAATTGATTTCTGCTTGCA ATTATCACTGATTAAAACTATTAATAATTTAACTAGTTGTATACAAGG
>mGR02 aa MESVLHNFATVLIYVEFIFGNLSNG FIVLSNFLDWVIKQKLSLIDKILLT LAISRITLIWEIYAWFKSLYDPSSF LIGIEFQIIYFSWVLSSHFSLWLAT TLSVFYLLRIANCSWQIFLYLKWRL KQLIVGMLLGSLVFLLGNLMQSMLE ERFYQYGRNTSVNTMSNDLAMWTEL IFFNMAMFSVIPFTLALISFLLLIF SLWKHLQKMQLISRRHRDPSTKAHM NALRIMVSFLLLYTMHFLSLLISWI AQKHQSELADIIGMITELMYPSVHS CILILGNSKLKQTSLCMLRHLRCRL KGENITIAYSNQITSFCVFCVANKS MR	>mGR02 nt CAGCACAGTGAAAAACTCATGGGCCACTTGGTCACCCAGGGACAGGCGACG CTGTTATATGCCAAGCTTTCTATGAACATGGAATCTGTCCTTCACAACTTT GCCACTGTACTAATATACGTGGAGTTTATTTTTGGGAATTTGAGCAATGGA TTCATAGTGTTGTCAAACTTCTTGGACTGGGTCATTAAACAAAAGCTTTCC TTAATAGATAAAATTCTTCTTACATTGGCAATTTCAAGAATCACTCTCATC TGGGAAATATATGCTTGGTTTAAAAGTTTATATGATCCATCTTCCTTTTTA ATTGGAATAGAATTTCAAATTATTTATTTTAGCTGGGTCCTTTCTAGTCAC TTCAGCCTCTGGCTTGCCACAACTCTCAGCGTCTTTTATTTACTCAGAATA GCTAACTGCTCCTGGCAGATCTTTCTCTATTTGAAATGGAGACTTAAACAA CTGATTGTGGGGATGTTGCTGGGAAGCTTGGTGTTCTTGCTTGGAAATCTG ATGCAAAGCATGCTTGAAGAGAGGTTCTATCAATATGGAAGGAACACAAGT GTGAATACCATGAGCAATGACCTTGCAATGTGGACCGAGCTGATCTTTTTC AACATGGCTATGTTCTCTGTAATACCATTTACATTGGCCTTGATTTCTTTT CTCCTGCTAATCTTCTCTTTGTGGAAACATCTCCAGAAGATGCAGCTCATT TCCAGAAGACACAGAGACCCTAGCACCAAGGCCCACATGAATGCCTTGAGA ATTATGGTGTCCTTCCTCTTGCTCTATACCATGCATTTCCTGTCTCTTCTT ATATCATGGATTGCTCAAAAGCATCAGAGTGAACTGGCTGATATTATTGGT ATGATAACTGAACTCATGTATCCTTCAGTCCATTCATGTATCCTGATTCTA GGAAATTCTAAATTAAAGCAGACTTCTCTTTGTATGCTGAGGCATTTGAGA TGTAGGCTGAAAGGAGAGAATATCACAATTGCATATAGCAACCAAATAACT AGCTTTTGTGTATTCTGTGTTGCAAACAAATCTATGAGGTAGTTGTTCAAG GAATCCTTCCTTGACTTATTGTATCATGGAAGTCATATGGGGGAGTCTGAA AGAGCTGTCTTCTGTAAGCAAGGTTTGTATACACTAGTGGGGCTGGGACAC CAACCCAAGCACAAAACCTAGCTATAACCTATCCTGGCTGCAGGATATGCT GGAACAATGGTGGCTTGGAAATTGTGGGACTGGCAAAGCAATAGCTAGTCT AACTTGAGGCCCATTCCACAGCAGGAAGCTCATGCCCACCTCTGCCTGGAT GGCCAGGAAGCAAAATCTTGATGGCCCCAAGACCTATGGTAAACTGAACAC TACTGGAAAAAGAAAGACTCGTGTTAATGATCTATCAAATATTTCCTAATG ATATTCTGATAAACTCATATATTAGTCCCTGTCCTAATCATCATCACTGGG ACTCCTTCCCAGCACCTGATGGGAGCAGATAGAGATCTACATCCAAATAGT AAGTGTATCTTGGGGAACTCCACTTAAGAATAGAAGGAACAATTATGAGAG CCAGAGTGATCCAGAACACTAGGATCACAGAATCAACTAAGCAGCATGCAT AGGGGTTAATGGAGACTGAAGTGGCAATCACAGAGCCTGCATAGGTCTACA CTAAGTCCTCTGTGTATATACTGTGGCTGTTTAGCTTAGGAATTTTGTTGG ACTCCTAACAATGGATAAGGAATTC
>mGR03 aa MVLTIRAILWVTLITIISLEFIIGI LGNVFIALVNIIDWVKRGKISAVDK TYMALAISRTAFLLSLITGFLVSLL DPALLGMRTMVRLLTISWMVTNHFS VWFATCLSIFYFLKIANFSNSIFLV LKWEAKKVVSVTLVVSVIILIMNII VINKFTDRLQVNTLQNCSTSNTLKD	>mGR03 nt CTTTAATAGCAGGGTGTGAATATTTAAATTTTCTTTCTGCAGCAACTACTG AGGGCTTCAGACTGCTGTATACAGGGCATGAAGCATCTGGATGAAGTTCAG CTGTGCTGCCTTTGACAACAATTTTTTGTGTATGTGTGGAGAACATAAACC ATTTCATTAGTGAAATTTGGCTTTTGGGTGACATTGTCTATGATAGTTCTG AAAGTGATTATGTTAAGAATCAGACACAGCCGTCTAGAAGATTGTATTAAC ACATCTTTGGTAGTTCAGAAGAAATTAGATCATCATGGTGTTGACAATAAG GGCTATTTTATGGGTAACATTGATAACTATTATAAGTCTGGAGTTTATCAT AGGAATTTTAGGAAATGTATTCATAGCTCTCGTGAACATCATAGACTGGGT

YGLFLFISTGFTLTPFA LTMFLL LIFSLWRHLKNMCHSATGSRDVSTV AHIKGLQTVVTFLLLYTAFVMSLLS ESLNINIQHTNLLSHFLRSIGVAFP TGHSCVLILGNSKLRQASLSVILWL RYKYKHIENWGP	TAAAAGAGGAAAGATCTCTGCAGTGGATAA CTATATGGCCCTGGCCAT CTCCAGGACTGCTTTTTTATTGTCACTAAT CAGGGTTCTTGGTATCATT ATTGGACCCAGCTTTATTGGGAATGAGAACGATGGTAAGGCTCCTTACTAT TTCCTGGATGGTGACCAATCATTTCAGTGTCTGGTTTGCAACATGCCTCAG TATCTTTTATTTTCTCAAGATAGCTAATTTCTCAAATTCTATTTTCCTTGT TCTCAAATGGGAAGCTAAAAAAGTGGTATCAGTGACATTGGTGGTATCTGT GATAATCTTGATCATGAACATTATAGTCATAAACAAATTCACTGACAGACT TCAAGTAAACACACTCCAGAACTGTAGTACAAGTAACACTTTAAAAGATTA TGGGCTCTTTTTATTCATTAGCACTGGGTTTACACTCACCCCATTCGCTGT GTCTTTGACAATGTTTCTTCTGCTCATCTTCTCCCTGTGGAGACATCTGAA GAATATGTGTCACAGTGCCACAGGCTCCAGAGATGTCAGCACAGTGGCCCA CATAAAAGGCTTGCAAACTGTGGTAACCTTCCTGTTACTATATACTGCTTT TGTTATGTCACTTCTTTCAGAGTCTTTGAATATTAACATTCAACATACAAA TCTTCTTTCTCATTTTTTACGGAGTATAGGAGTAGCTTTTCCCACAGGCCA CTCCTGTGTACTGATTCTTGGAAACAGTAAGCTGAGGCAAGCCTCTCTTTC TGTGATATTGTGGCTGAGGTATAAGTACAAACATATAGAGAATTGGGGCCC CTAAATCATATCAGGGATCCTTTTCCACATTCTAGAAAAAAATCAGTTAAT AAGAACAGGAATTTAGGAAGGAATCTGAAATTATGAATCTCATAGGCCATG AACCTTCAGACAAAGGATTCATTAGAGAGATAGAGAGAGAACATTGTTATC TGTAACTCGACAGGCAACACTGTAGATTATGAAAATAAATGTCAGTCTGTA ATGGAAAGCAAAACATGCTATATTTTATTAATTGGTTTTGGTTTAAGGTCG GGATA
>mGR04 aa MLSALESILLSVATSEAMLGVLGNT FIVLVNYTDWVRNKKLSKINFILTG LAISRIFTIWIITLDAYTKVFLLTM LMPSSLHECMSYIWVIINHLSVWFS TSLGIFYFLKIANFSHYIFLWMKRR ADKVFVFLIVFLIITWLASFPLAVK VIKDVKIYQSNTSWLIHLEKSELLI NYVFANMGPISLFIVAIIACFLLTI SLWRHSRQMQSIGSGFRDLNTEAHM KAMKVLIAFIILFILYFLGILIETL CLFLTNNKLLFIFGFTLSAMYPCCH SFILILTSRELKQDTMRALQRLKCC ET	>mGR04 nt CTGCAGCAGGTAAATCACACCAGATCCAGCAGAAGCCTTCTTGGAAATTGG CAGAGATGCTGAGTGCACTGGAAAGCATCCTCCTTTCTGTTGCCACTAGTG AAGCCATGCTGGGAGTTTTAGGGAACACATTTATTGTACTTGTAAACTACA CAGACTGGGTCAGGAATAAGAAACTCTCTAAGATTAACTTTATTCTCACTG GCTTAGCAATTTCCAGGATTTTTACCATATGGATAATAACTTTAGATGCAT ATACAAAGGTTTTCCTTCTGACTATGCTTATGCCGAGCAGTCTACATGAAT GCATGAGTTACATATGGGTAATTATTAACCATCTGAGCGTTTGGTTTAGCA CCAGCCTCGGCATCTTTTATTTTCTGAAGATAGCAAATTTTTCCCACTACA TATTTCTCTGGATGAAGAGAAGAGCTGATAAAGTTTTTGTCTTTCTAATTG TATTCTTAATTATAACGTGGCTAGCTTCCTTTCCGCTAGCTGTGAAGGTCA TTAAAGATGTTAAAATATATCAGAGCAACACATCCTGGCTGATCCACCTGG AGAAGAGTGAGTTACTTATAAACTATGTTTTTGCCAATATGGGGCCCATTT CCCTCTTTATTGTAGCCATAATTGCTTGTTTCTTGTTAACCATTTCCCTTT GGAGACACAGCAGGCAGATGCAATCCATTGGATCAGGATTCAGAGATCTCA ACACAGAAGCTCACATGAAAGCCATGAAAGTTTTAATTGCATTTATCATCC TCTTTATCTTATATTTTTTGGGTATTCTCATAGAAACATTATGCTTGTTTC TTACAAACAATAAACTTCTCTTTATTTTTGGCTTCACTTTGTCAGCCATGT ATCCCTGTTGCCATTCCTTTATCCTAATTCTAACAAGCAGGGAGCTGAAGC AAGACACTATGAGGGCACTGCAGAGATTAAAATGCTGTGAGACTTGACAGA GAAATGAATGTTCTGGCACAGTTCAGCAGGGAATCCCTGGAGCCCTTTCCA TTCCCACTATGTTCTCACACTGTCTTTAGTTGAATTGTTAAAAGTTTTTGA AACCTTTGGCAACTGATTGACTGCAGCTACGCCAGTGTAAGATTTTCATAG TAAGAGCAAACATTGAAAATAAGACTTCTCAGTCTTATTTCATTGAGTTTC TAAAGCATTGACACCCATTCACCAGAAAAACCAAAGGGGAAGAGAGGAGTT TTCAGACATGTGTGATGAATCTTGATATTTAGGACATGGAATTGAGGAG-C CAGAGGGATGCTACCGTGTGTCTACAGCTTTGTTTGTTAAATAGCTACTTT TCCTTTCCCAGTTAGTTAAAGTAGATGCTTGGAGTAGTGGTGAAAATCATG GCAGTAGATGGGATCTGTGGGAAGTGGTTGAGGAAGCAGGCTGTTTCTGAA CGAAGAGACCAGAGGACTGATTGAACTGGTCATTGTGTATATCAAAAATAG TGATTTCAGATGAAGCCAAGTTGTAGAGCAAAGATATCTGAGGAAGAATTC
>mGR05 aa MLSAAEGILLSIATVEAGLGVLGNT FIALVNCMDWAKNNKLSMTGFLLIG LATSRIFIVWLLTLDAYAKLFYPSK YFSSSLIEIISYIWMTVNHLTVWFA TSLSIFYFLKIANFSDCVFLWLKRR TDKAFVFLLGCLLTSWVISFSFVVK VMKDGKVNHRNRTSEMYWEKRQFTI NYVFLNIGVISLFMMTLTACFLLIM	>mGR05 nt ATGCTGAGTGCGGCAGAAGGCATCCTCCTTTCCATTGCAACTGTTGAAGCT GGGCTGGGAGTTTTAGGGAACACATTTATTGCACTGGTAAACTGCATGGAC TGGGCCAAGAACAATAAGCTTTCTATGACTGGCTTCCTTCTCATCGGCTTA GCAACTTCCAGGATTTTTATTGTGTGGCTATTAACTTTAGATGCATATGCA AAGCTATTCTATCCAAGTAAGTATTTTTCTAGTAGTCTGATTGAAATCATC TCTTATATATGGATGACTGTGAATCACCTGACTGTCTGGTTTGCCACCAGC CTAAGCATCTTCTATTTCCTGAAGATAGCCAATTTTTCCGACTGTGTATTT CTCTGGTTGAAGAGGAGAACTGATAAAGCTTTTGTTTTTCTCTTGGGGTGT TTGCTAACTTCATGGGTAATCTCCTTCTCATTTGTTGTGAAGGTGATGAAG

SLWRHSRQMQSGVSGF[redacted]NTEAHV KAIKFLISFIILFVLYFIGVSIEII CIFIPENKLLFIFGFTTASIYPCCH SFILILSNSQLKQAFVKVLQGLKFF	GACGGTAAAGTGAATCATAGAAACAGGAC[redacted]GGAGATGTACTGGGAGAAA AGGCAATTCACTATTAACTACGTTTTCCT[redacted]ATATTGGAGTCATTTCTCTC TTTATGATGACCTTAACTGCATGTTTCTTGTTAATTATGTCACTTTGGAGA CACAGCAGGCAGATGCAGTCTGGTGTTTCAGGATTCAGAGACCTCAACACA GAAGCTCATGTGAAAGCCATAAAATTTTTAATTTCATTTATCATCCTTTTC GTCTTGTATTTTATAGGTGTTTCAATAGAAATTATCTGCATATTTATACCA GAAAACAAACTGCTATTTATTTTTGGTTTCACAACTGCATCCATATATCCT TGCTGTCACTCATTTATTCTAATTCTATCTAACAGCCAGCTAAAGCAAGCC TTTGTAAAGGTACTGCAAGGATTAAAGTTCTTTTAG
>mGR06 aa MLTVAEGILLCFVTSGSVLGVLGNG FILHANYINCVRKKFSTAGFILTGL AICRIFVICIIISDGYLKLFSPHMV ASDAHIIVISYIWVIINHTSIWFAT SLNLFYLLKIANFSHYIFFCLKRRI NTVFIFLLGCLFISWSIAFPQTVKI FNVKKQHRNVSWQVYLYKNEFIVSH ILLNLGVIFFFMVAIITCFLLIISL WKHNRKMQLYASRFKSLNTEVHVKV MKVLISFIILLILHFIGILIETLSF LKYENKLLLILGLIISCMYPCCHSF ILILANSQLKQASLKALKQLKCHKK DKDVRVTW	>mGR06 nt TATAGTTGCAGCAGAAGCAACGTTAGGGATCTGTAGAGATGCTGACTGTA GCAGAAGGAATCCTCCTTTGTTTTGTAACTAGTGGTTCAGTCCTGGGAGT TCTAGGAAATGGATTTATCCTGCATGCAAACTACATTAACTGTGTCAGAA AGAAGTTCTCCACAGCTGGCTTTATTCTCACAGGCTTGGCTATTTGCAGA ATCTTTGTCATATGTATAATAATCTCTGATGGATATTTAAAATTGTTTTC TCCACATATGGTTGCCTCTGATGCCCACATTATAGTGATTTCTTACATAT GGGTAATTATCAATCATACAAGTATATGGTTTGCCACCAGCCTCAACCTC TTCTATCTCCTGAAGATAGCAAATTTTTCTCACTACATCTTCTTCTGCTT GAAGAGAAGAATCAATACAGTATTTATCTTTCTCCTGGGATGCTTATTTA TATCATGGTCAATTGCTTTCCCACAAACAGTGAAGATATTTAATGTTAAA AAGCAGCACAGAAaTGTTTCCTGGCAGGTTTACCTCTATAAGAATGAGTT CATtGTAAGCCACATTCTTCTCAACCTGGGAGTTATATTCTTCTTTATGG TGGCTATCATTACATGCTTCCTATTAATTATTTCACTTTGGAAACATAAC AGAAAGATGCAGTTGTATGCCTCAAGATTCAAAAGCCTTAACACAGAAGT ACATGTGAAAGTCATGAAAGTTTTAATTTCTTTTATTATCCTGTTAATCT TGCATTTCATAGGGATTTTGATAGAAACATTGAGCTTTTTAAAATATGAA AATAAACTGCTACTTATTTTGGGTTTGATAATTTCATGCATGTATCCTTG CTGTCATTCATTTATCCTAATTCTAGCAAACAGTCAGCTGAAGCAGGCTT CTTTGAAGGCACTGAAGCAATTAAAATGCCATAAGAAAGACAAGGACGTC AgAGTGACATGGTAGACTTATGGAGAAATGAATGGTCACAAGAAATAGCC TGGTGTGGAGATGTTGATATCTCTAAAGACCGTTTCACTTCCAAATTCTT GCAATTATTTAAAAAAAAAAAGTCTTGCTGATATCATGGAATCATGGGAAA TGTTGCAATTGTGTTTTGGGGACAGGGTGACCAGTGAAGGTATGGTTAAG CAGCGAAACACTCATACAGCTCGTTCGTTCTTTTTGTATTTTATTTTGTG TTGGTGGCCTTCCAAGACATGATTTCTCTATGTAAGTTTTGG
>mGR07 aa MLNSAEGILLCVVTSEAVLGVLGDT YIALFNCMDYAKNKKLSKIGFILIG LAISRIGVVWIIILQGYIQVFFPHM LTSGNITEYITYIWVFLNHLSVWFV TNLNILYFLKIANFSNSVFLWLKRR VNAVFIFLSGCLLTSWLLCFPQMTK ILQNSKMHQRNTSWVHQRKNYFLIN QSVTNLGIFFFIIVSLITCFLLIVF LWRHVRQMHSDVSGFRDHSTKVHVK AMKFLISFMVFFILHFVGLSIEVLC FILPQNKLLFITGLTATCLYPCGHS IIVILGNKQLKQASLKALQQLKCCE TKGNFRVK	>mGR07 nt TTCATAATGAAGAGGAGGCAGGGCAATGTTGGTTTCTGTTGTCTGACCAGT GTATTTGACAGTGATACTACACATTTGATTGCTAAATGCAAATAGTTCCAA AGGAACAAGTAAATTTTATGAAATAGAAGCTTCTATTTGCTTATTAACAAA CTGCAAGCAAACATTAGTCTGCACACATTTTATAGACAAGCTAAATCTTCA AAAGCAATAAAAAAGAGCACCCATAAAGTTCTGACTCTATCACATGACAAT AGGCTTGAAAAGATTGTCTATGTAGATAAAGAAGATGGCATAACTTCTCCA TCAAGAAGCCAGTATATGGGACATTCTCCAGCAGATAATTTACAATAGATG CAGCAGAAGTAACCTTAGAGATCTGTAAAGATGCTGAATTCAGCAGAAGGC ATCCTCCTTTGTGTTGTCACTAGTGAGGCTGTGCTCGGAGTTTTAGGGGAC ACATATATTGCACTTTTTAACTGCATGGACTATGCTAAGAACAAGAAGCTC TCTAAGATCGGTTTCATTCTCATTGGCTTGGCGATTTCCAGAATTGGTGTT GTATGGATAATAATTTTACAAGGGTATATACAAGTATTTTTTCCACACATG CTTACCTCTGGAAACATAACTGAATATATTACTTACATATGGGTATTTCTC AATCACTTAAGTGTCTGGTTTGTCACCAACCTCAACATCCTCTACTTTCTA AAGATAGCTAATTTTTCCAACTCTGTATTTCTCTGGCTGAAAAGGAGAGTC AATGCAGTTTTTATCTTTCTGTCAGGATGCTTACTTACCTCATGGTTACTA TGTTTTCCACAAATGACAAAGATACTTCAAAATAGTAAAATGCACCAGAGA AACACATCTTGGGTCCACCAGCGGAAAAATTACTTTCTTATTAACCAAAGT GTGACCAATCTGGGAATCTTTTTCTTCATTATTGTATCCCTGATTACCTGC TTTCTGTTGATTGTTTTCCTCTGGAGACATGTCAGACAAATGCACTCAGAT GTTTCAGGATTCAGAGACCACAGCACAAAAGTACATGTGAAAGCTATGAAA TTTCTAATATCTTTTATGGTCTTCTTTATTCTGCATTTTGTAGGCCTTTCC ATAGAAGTGCTATGCTTTATTCTGCCACAAAATAAACTGCTCTTTATAACT GGTTTGACAGCCACATGCCTCTATCCCTGCGGTCACTCAATCATCGTAATT TTAGGAAATAAGCAGTTAAAGCAAGCCTCTTTGAAGGCACTGCAGCAACTA AAATGCTGTGAGACAAAAGGAAATTTCAGAGTCAAATAAATGGGTTTGCAA ATAAATAGCTGCCTTGTTCTTcCACTGGTTTTTACCCTGTTAGTTGATGTT

	ATGAAAAGTTCCTGCTATGGTTGATGACA[redacted]CAAGGAATCTATTTTTCTG GTGGCATGTTAAGTCCACGTGAAGCCTCA[redacted]CATACTGTGACTTGACTAT GCAAATTCTTTCCACAAAATAACCAGATAACATTCAGCCTGGAGATAAATT CATTTAAAGGCTTTTATGGTGAGGATAAACAAAAAAAAAAAATCATTTTTC TGTGATTCACTGTAACTCCCAGGATGAGTAAAAGAAAACAAGACAAATGGT TGTGATCAGCCTTTGTGTGTCTAGACAGAGCTAGGGACCAGATGTTGATGC TTGTGTGTGGTTTTGAGTTCTTTAAGAAGTTATTGCCTCTCTGCCATTCGG TATTCCTCAGGTGAGAATTC
>mGR08 aa MLWELYVFVFAASVFLNFVGIIANL FIIVIIIKTWVNSRRIASPDRILFS LAITRFLTLGLFLLNSVYIATNTGR SVYFSTFFLLCWKFLDANSLWLVTI LNSLYCVKITNFQHPVFLLLKRTIS MKTTSLLLACLLISALTTLLYYMLS QISRFPEHIIGRNDTSFDLSDGILT LVASLVLNSLLQFMLNVTFASLLIH SLRRHIQKMQRNRTSFWNPQTEAHM GAMRLMICFLVLYIPYSIATLLYLP SYMRKNLRAQAICMIITAAYPPGHS VLLIITHHKLKAKAKKIFCFYK	>mGR08 nt AAGCTTGTTTGTAATTAGGCATTCCTAAGAAAATAAGAACAGGAGTGAAGA AATAGTAATTTAATCCTTGAAAGATTTGCATCTCAGTAAAAGCAGCTGCCT CTTAGACCAGAAATGGTGTTTGCCATGCTGGAAAATAAAAAGGAGACCTCT TTCCAGGCTGCATCCTGTGTCTGCTTACTTATTTCAGTTTGTTTTCATCGG CACCAAACGAGGAAAGATGCTCTGGGAACTGTATGTATTTGTGTTTGCTGC CTCGGTTTTTTTAAATTTTGTAGGAATCATTGCAAATCTATTTATTATAGT GATAATTATTAAGACTTGGGTCAACAGTCGCAGAATTGCCTCTCCGGATAG GATCCTGTTCAGCTTGGCCATCACTAGATTCCTGACTTTGGGGTTGTTTCT ACTGAACAGTGTCTACATTGCTACAAATACTGGAAGGTCAGTCTACTTTTC CACATTTTTTCTATTGTGTTGGAAGTTTCTGGATGCAAACAGTCTCTGGTT AGTGACCATTCTGAACAGCTTGTATTGTGTGAAGATTACTAATTTTCAACA CCCAGTGTTTCTCCTGTTGAAACGGACTATCTCTATGAAGACCACCAGCCT GCTGTTGGCCTGTCTTCTGATTTCAGCCCTCACCACTCTCCTATATTATAT GCTCTCACAGATATCACGTTTTCCTGAACACATAATTGGGAGAAATGACAC GTCATTTGACCTCAGTGATGGTATCTTGACGTTAGTAGCCTCTTTGGTCCT GAACTCACTTCTACAGTTTATGCTCAATGTGACTTTTGCTTCCTTGTTAAT ACATTCCTTGAGAAGACATATACAGAAGATGCAGAGAAACAGGACCAGCTT TTGGAATCCCCAGACGGAGGCTCACATGGGTGCTATGAGGCTGATGATCTG TTTCCTCGTGCTCTACATTCCATATTCAATTGCTACCCTGCTCTATCTTCC TTCCTATATGAGGAAGAATCTGAGAGCCCAGGCCATTTGCATGATTATTAC TGCTGCTTACCCTCCAGGACATTCTGTCCTCCTCATTATCACACATCATAA ACTGAAAGCTAAAGCAAAGAAGATTTTCTGTTTCTACAAGTAGCAGAATTT CATTAGTAGTTAACAGCATCAATTCATGGTTTGGTTGCATTAGAAATGTCT CAGTGATCTAAGGACTTAATTTTGTGATCTTGTATCTGGCATCCTGACCCT GAGACTAAGTGCTTATATTTTGGTCAATACAGCATCTTTTGGCTAATATTT TAAAGTAAATCACATTCCATAAGAAATTGTTTAAGGGATTTACGTATTTTT CATGGCTATCACATTCCTAGACAATGGAAATCACCATACTGTTTCGCTAGC TACTGAAGTACCAGGGGAAAGTCCATGAATGAAGGCCACATTGTGATGTTC TTGGTTAGCACAGATTAGAGAATTTGGCCTCAACTGAGCAAGATATC
>mGR09 aa MEHLLKRTFDITENILLIILFIELI IGLIGNGFTALVHCMDWVKRKKMSL VNKILTALATSRIFLLWFMLVGFPI SSLYPYLVTTRLMIQFTSTLWTIAN HISVWFATCLSVFYFLKIANFSNSP FLYLKRRVEKVVSVTLLVSLVLLFL NILLLNLEINMCINEYHQINISYIF ISYYHLSCQIQVLGSHIIFLSVPVV LSLSTFLLLIFSLWTLHKRMQQHVQ GGRDARTTAHFKALQAVIAFLLLYS IFILSLLLQFWIHGLRKKPPFIAFC QVVDTAFPSFHSYVLILRDRKLRHA SLSVLSWLKCRPNYVK	>mGR09 nt GAATTCAGAAATCATCAAAAAATCTTCAAAACTACATGTTTAAAATAGCAC TTCAAATGAATACATTTGCAAATCTTTACAACTAATACATAAAATGGAGCA TCTTTTGAAGAGAACATTTGATATCACCGAGAACATACTTCTAATTATTTT ATTCATTGAATTAATAATTGGACTTATAGGAAACGGATTCACAGCCTTGGT GCACTGCATGGACTGGGTTAAGAGAAAAAAAATGTCATTAGTTAATAAAAT CCTCACCGCTTTGGCAACTTCTAGAATTTTCCTGCTCTGGTTCATGCTAGT AGGTTTTCCAATTAGCTCACTGTACCCATATTTAGTTACTACTAGACTGAT GATACAGTTCACTAGTACTCTATGGACTATAGCTAACCATATTAGTGTCTG GTTTGCTACATGCCTCAGTGTCTTTTATTTTCTCAAGATAGCCAATTTTTC TAATTCTCCTTTTCTCTATCTAAAGAGGAGAGTTGAAAAAGTAGTTTCAGT TACATTACTGGTGTCTCTGGTCCTCTTGTTTTTAAATATTTTACTACTTAA TTTGGAAATTAACATGTGTATAAATGAATATCATCAAATAAACATATCATA CATCTTCATTTCTTATTACCATTTAAGTTGTCAAATTCAGGTGTTAGGAAG TCACATTATTTTCCTGTCTGTCCCCGTTGTTTTGTCCCTGTCAACTTTTCT CCTGCTCATCTTCTCCCTGTGGACACTTCACAAGAGGATGCAGCAGCATGT TCAGGGAGGCAGAGATGCCAGAACCACGGCCCACTTCAAAGCCTTGCAAGC AGTGATTGCCTTTCTCCTACTATACTCCATTTTTATCCTGTCACTGTTACT ACAATTTTGGATCCATGGATTAAGGAAGAAACCTCCTTTCATTGCATTTTG TCAGGTTGTAGATACAGCTTTTCCTTCATTCCATTCATATGTCTTGATTCT GAGAGACAGGAAGCTGAGACACGCCTCTCTCTCTGTGTTGTCGTGGCTGAA ATGCAGGCCAAATTATGTGAAATAATATTTCTTTGTATTTTCATTTTCAAT TTTAAAATATTCTTAGAATTTGACTGCATGTATTTCATCTTTTATTTGAAA CAACCACTAATTAAAGCTATTACTAATTTAGCAAGTCGTATACAAGGTTAT TTTTTAATACACATATCAAAAACTGACATGTTTATGTTCTACAAAAACCTG

	AATATATCAAAATTATATAAATTTTGTA[redacted]ACGATTAACAATGGAGTTTT TTTATTTATGACCTGTCACGGGACTCCGGTGGAGTCAGCTTGTCAGATGAA AGTCTGAAAGCTT
>mGR10 aa MFSQIISTSDIFTFTIILFVELVIG ILGNGFIALVNIMDWTKRRSISSAD QILTALAITRFLYVWFMIICILLFM LCPHLLTRSEIVTSIGIIWIVNNHF SVWLATCLGVFYFLKIANFSNSLFL YLKWRVKKVVLMIIQVSMIFLILNL LSLSMYDQFSIDVYEGNTSYNLGDS TPFPTISLFINSSKVFVITNSSHIF LPINSLFMLIPFTVSLVAFLMLIFS LWKHHKKMQVNAKPPRDASTMAHIK ALQTGFSFLLLYAVYLLFIVIGMLS LRLIGGKLILLFDHISGIGFPISHS FVLILGNNKLRQASLSVLHCLRCRS KDMDTMGP	>mGR10 nt GAATTCAACATCTTATTCAACTTCAGAAAACTGGATATTAGACACAGTGTC TGGATGAAGCAGAGGTGATCTCTTTGGGAAAAAAAGCCAAGTAGTCATAAA GAATTTATGAAACAATTCCTGGGATTGTTTATATTTGTTACAAACAAATTT ATATGTTTGTTAGTCAGTAATGTATAAGTGGGATTTTAAAGCATGATTATC TTGAATTTTTAACAAAAAACATGTAGTGCTTTTTAAATGTAGCAGAAACAT TAAAAATTGAAGCATGTTCTCACAGATAATAAGCACCAGTGATATTTTTAC TTTTACAATAATATTATTTGTGGAATTAGTAATAGGAATTTTAGGAAATGG ATTCATAGCACTAGTGAATATCATGGACTGGACCAAGAGAAGAAGCATTTC ATCAGCGGATCAGATTCTCACTGCTTTGGCCATTACCAGATTTCTCTATGT GTGGTTTATGATCATTTGTATATTGTTATTCATGCTGTGCCCACATTTGCT TACAAGATCAGAAATAGTAACATCAATTGGTATTATTTGGATAGTGAATAA CCATTTCAGCGTTTGGCTTGCCACATGCCTCGGTGTCTTTTATTTTCTGAA GATAGCCAATTTTTCTAACTCTTTGTTTCTTTACCTAAAGTGGAGAGTTAA AAAAGTAGTTTTAATGATAATACAGGTATCAATGATTTTCTTGATTTTAAA CCTGTTATCTCTAAGCATGTATGATCAGTTCTCAATTGATGTTTATGAAGG AAATACATCTTATAATTTAGGGGATTCAACCCCATTTCCCACAATTTCCTT ATTCATCAATTCATCAAAAGTTTTCGTAATCACCAACTCATCCCATATTTT CTTACCCATCAACTCCCTGTTCATGCTCATACCCTTCACAGTGTCCCTGGT AGCCTTTCTCATGCTCATCTTCTCACTGTGGAAGCATCACAAAAAGATGCA GGTCAATGCCAAACCACCTAGAGATGCCAGCACCATGGCCCACATTAAAGC CTTGCAAACAGGGTTCTCCTTCCTGCTGCTGTATGCAGTATACTTACTTTT TATTGTCATAGGAATGTTGAGCCTTAGGTTGATAGGAGGAAAATTAATACT TTTATTTGACCACATTTCTGGAATAGGTTTTCCTATAAGCCACTCATTTGT GCTGATTCTGGGAAATAACAAGCTGAGACAAGCCAGTCTTTCAGTGTTGCA TTGTCTGAGGTGCCGATCCAAAGATATGGACACCATGGGTCCATAAAAAAT TTCAGAGGTCATTGGGAAACATTTTGAGATCTTATAGGGGAAAAAGAAAAT GTGGGGCTTCAAAGCTGGTAGGAGTAATATAGAGAAGGATAGGAG
>mGR11 aa (notional!) MEHPLRRTFDFSQSILLTILFIELI IGLIRNGLMVLVHCIDWVKRKKFHL LIKSSPLWQTSRICLLWFMLIHLLI TLLYADLASTRTMMQFASNPWTISN HISIWLATCLGVFYFLKIANFSNST FLYLKWRVQFLLLNILLVKFEINMW INEYHQINIPYSFISYYQXCQIQVL SLHIIFLSVPFILSLSTFLLLIFSL WTLHQRMQQHVQGYRDASTMAHFKA LQAVIAFLLIHSIFILSLLLQLWKH ELRKKPPFVVFCQVAYIAFPSSHSY VFILGDRKLRQACLSVLWRLKCRPN YVG	>mGR11 nt AATAATGTATGTGGAAGAGTTAAGTATAAATGTTGTATGAGAATGAACTCA GAAATCATCAAAAATCTTTAAAACTGCATGTTAAAAATCACACTTCAAATG AATATATTTGTAATTCTTTAGAACTAATAAATAAAATGGAGCATCCTTTGA GGAGAACATTTGATTTCTCCCAGAGCATACTTCTAACCATTTTATTCATTG AATTAATAATTGGACTTATAAGAAATGGATTAATGGTATTGGTGCACTGCA TAGATTGGGTTAAGAGAAAAAAATTTCATTTGTTAATCAAATCCTCACCAC TTTGGCAAACTTCCAGAATTTGTCTGCTCTGGTTCATGCTAATACATCTCC TGATTACTTTATTGTATGCAGATTTAGCTAGTACTAGAACGATGATGCAAT TCGCTAGCAATCCATGGACTATATCTAACCATATCAGCATCTGGCTTGCTA CATGCCTTGGTGTCTTTTATTTTCTCAAGATAGCCAATTTTTCTAACTCTA CTTTTCTCTATCTAAAATGGCGAGTTCAGTTCCTCTTGTTAAATATTTTAC TGGTTAAATTTGAGATTAACATGTGGATAAATGAATATCATCAAATAAACA TACCATACAGCTTCATTTCTTATTACCAAATTGTCAAATACAGGTGTTAAG TCTTCACATTATTTTCCTGTCTGTCCCCTTTATTTTGTCCCTGTCAACTTT TCTCCTGCTCATCTTCTCCCTGTGGACACTTCACCAGAGGATGCAGCAGCA TGTTCAAGGATACAGAGATGCCAGCACAATGGCCCACTTCAAAGCCTTGCA AGCAGTGATTGCCTTTCTCTTAATACACTCCATTTTTATCCTGTCACTGTT ACTACAACTTTGGAAACATGAATTAAGGAAGAAACCTCCTTTTGTTGTATT TTGTCAGGTTGCATATATAGCTTTTCCTTCATCCCATTCATATGTCTTCAT TCTGGGAGACAGAAAGCTGAGACAGGCTTGTCTCTCTGTGTTGTGGAGGCT GAAATGCAGGCCAAATTATGTGGGATAAAATCTCTTTGTGCTTTCATTTCC AATTCTTAAATATTCTTTGATTTTGACTGCATAAATT
>mGR12 aa (partial) GAIVNVDFLIGNVGNGFIVVANIMD LVKRRKLSSVDQLLTALAVSRITLL WYLYIMKRTFLVDPNIGAIMQSTRL TNVIWIISNHFSIWLATTLSIFYFL KIANFSNSIFCYLRWRFEKVILMAL LVSLVLLFIDILVTNMYINIWTDEF	>mGR12 nt (truncated) TTTTCAGCAGTGACTTTGGGAAGCAGAACGTCCTCTTAGAGACAGTGGGTG CTGCTATCCTAGTTAATGTGGAGCAATAGTTAATGTGGATTTCCTAATTGG AAATGTTGGGAATGGATTCATTGTTGTGGCAAACATAATGGACTTGGTCAA GAGAAGAAAGCTTTCTTCAGTGGATCAGCTGCTCACTGCACTGGCCGTCTC CAGAATCACTTTGCTGTGGTACCTGTACATAATGAAACGAACATTTTTAGT GGATCCAAACATTGGTGCAATTATGCAATCAACAAGACTGACTAATGTTAT CTGGATAATTTCTAACCATTTTAGTATATGGCTGGCCACCACCCTCAGCAT CTTTTATTTTCTCAAGATAGCAAATTTTTCTAACTCTATTTTCTGTTACCT

	GAGGTGGAGATTTGAAAAGGTGATTTTC[redacted]GCATTGCTGGTGTCCCTGGT CCTCTTGTTTATAGATATTTTAGTAACA[redacted]CATGTACATTAATATTTGGAC TGATGAATTC
>mGR13 aa MVAVLQSTLPIIFSMEFIMGTLGNG FIFLIVCIDWVQRRKISLVDQIRTA LAISRIALIWLIFLDWWVSVHYPAL HETGKMLSTYLISWTVINHCNFWLT ANLSILYFLKIANFSNIIFLYLKFR SKNVVLVTLLVSLFFLFLNTVIIKI FSDVCFDSVQRNVSQIFIMYNHEQI CKFLSFTNPMFTFIPFVMSTVMFSL LIFSLWRHLKNMQHTAKGCRDISTT VHIRALQTIIVSVVLYTIFFLSFFV KVWSFVSPERYLIFLFVWALGNAVF SAHPFVMILVNRRLRLASLSLIFWL WYRFKNIEV	>mGR13 nt AAGCTTGTTTGTGTTTGGATGAATTCTATTTATGTCTATCAATTTAAGATT TTCATATGAATCATTAAGAAATCTTGATAGTTGTTTGTGAGATATCACTTC TGCAATTTTTAAATGAAATTACACTCATATTTTGAAGGAACAATATGTTTT AAAGGAATATATTAACAAATCTTCAGCAGTTACCTCAGAAGTTTGGGTATT GTTTTACAGAAAATGGTGGCAGTTCTACAGAGCACACTTCCAATAATTTTC AGTATGGAATTCATAATGGGAACCTTAGGAAATGGATTCATTTTTCTGATA GTCTGCATAGACTGGGTCCAAAGAAGAAAAATCTCTTTAGTGGATCAAATC CGCACTGCTCTGGCAATTAGCAGAATCGCTCTAATTTGGTTGATATTCCTA GATTGGTGGGTGTCTGTTCATTACCCAGCATTACATGAAACTGGTAAGATG TTATCAACATATTTGATTTCCTGGACGGTGATCAATCATTGTAACTTTTGG CTTACTGCAAACTTGAGCATCCTTTATTTTCTCAAGATAGCCAACTTTTCT AACATTATTTTTCTTTATCTAAAGTTTAGATCTAAAAATGTGGTATTAGTG ACCCTGTTAGTGTCTCTATTTTTCTTGTTCTTAAATACTGTAATTATAAAA ATATTTTCTGATGTGTGTTTTGATAGTGTTCAAAGAAATGTGTCTCAAATT TTCATAATGTATAACCATGAACAAATTTGTAAATTTCTTTCCTTTACTAAC CCTATGTTCACATTCATACCTTTTGTTATGTCCACGGTAATGTTTTCTTTG CTCATCTTCTCCCTGTGGAGACATCTGAAGAATATGCAGCACACCGCCAAA GGATGCAGAGACATCAGCACCACAGTGCACATCAGAGCCCTGCAAACCATC ATTGTGTCTGTAGTGCTATACACTATTTTTTTTCTATCATTTTTTGTTAAA GTTTGGAGTTTTGTGTCACCAGAGAGATACCTGATCTTTTTGTTTGTCTGG GCTCTGGGAAATGCTGTTTTTTCTGCTCACCCATTTGTCATGATTTTGGTA AACAGAAGATTGAGATTGGCTTCTCTCTCTCTGATTTTTTGGCTCTGGTAC AGGTTTAAAAATATAGAAGTATAGGGTCCAAAGACCACCAAGGAATCATTT TCCTTATCCTAAAGAAAAATCAGGAG
>mGR14 aa MLSTMEGVLLSVSTSEAVLGIVGNT FIALVNCMDYNRNKKLSNIGFILTG LAISRICLVLILITEAYIKIFYPQL LSPVNIIELISYLWIIICQLNVWFA TSLSIFYFLKIANFSHYIFVWLKRR IDLVFFFLIGCLLISWLFSFPVVAK MVKDNKMLYINTSWQIHMKKSELII NYVFTNGGVFLFFMIMLIVCFLLII SLWRHRRQMESNKLGFRDLNTEVHV RTIKVLLSFIILFILHFMGITINVI CLLIPESNLLFMFGLTTAFIYPGCH SLILILANSRLKQCSVMILQLLKCC ENGKELRDT	>mGR14 nt CTGCAGGTATATACCTACCCTGAAGGCTTCATCTAGAGTAAACAAAGTAGT CTGTATAGTCTGCCATTCCTCAGATTCTCCTCAACTTCCCACCCTCCAGTG ACCTTTCTCCTTTTCTACAGTCAAACTATGGACCTCACAACCTGACACTTC TTCAGATGCAAAATATTCTCACAGAGACAAGTAAAACATACAAAACAAATA CTTTAATTTGCCTATTAACAAATGGCAAGAAAAGATTCAGGCTTGAACATC CTGTAGACAAGCTAAGGACAGGAGCAACTGAAGGGATCTCCATGAAGACCT TTCAGATTTCTACCAAAAGTAATTTTTAACTATATTTAAGTCTTTAAAGAA AGAAAGTAAAGCCACTCTTTTATTGAACAGCAATAGATTGGAATCTTAAAC AACTGCAACAGAAGCCATTTTAAAGATCAACAAAGATGCTGAGCACAATGG AAGGTGTCCTCCTTTCAGTTTCAACTAGTGAGGCTGTGCTGGGCATTGTAG GGAACACATTCATTGCACTTGTAAACTGTATGGACTATAACAGGAACAAGA AGCTCTCTAATATTGGCTTTATTCTCACTGGCTTGGCAATTTCCAGAATTT GCCTTGTGTTGATCTTAATCACAGAGGCATACATAAAAATATTCTATCCAC AGTTGCTGTCTCCTGTCAACATAATTGAGCTCATCAGTTATCTATGGATAA TTATCTGTCAATTGAATGTCTGGTTTGCCACTAGTCTCAGTATTTTTTATT TCCTGAAGATAGCAAATTTTTCCCACTACATATTTGTCTGGTTAAAAAGAA GAATTGATTTAGTTTTTTTCTTCCTGATAGGGTGCTTGCTTATCTCATGGC TATTTTCTTTCCCAGTTGTTGCGAAGATGGTTAAAGATAATAAAATGCTGT ATATAAACACATCTTGGCAGATCCACATGAAGAAAAGTGAGTTAATCATTA ACTATGTTTTCACCAATGGGGGAGTATTTTTATTTTTTATGATAATGTTAA TTGTATGTTTCCTGTTAATCATTTCACTTTGGAGACATCGCAGGCAGATGG AATCAAATAAATTAGGATTCAGAGATCTCAACACAGAAGTTCATGTGAGAA CAATAAAAGTTTTATTGTCTTTTATTATCCTTTTTATATTGCATTTCATGG GTATTACCATAAATGTAATTTGTCTGTTAATCCCAGAAAGCAACTTGTTAT TCATGTTTGGTTTGACAACTGCATTCATCTATCCCGGCTGCCACTCACTTA TCCTAATTCTAGCAAACAGTCGGCTGAAGCAGTGCTCTGTAATGATACTGC AACTATTAAAGTGCTGTGAGAATGGTAAAGAACTCAGAGACACATGACAGT CTGGAACACATGCAATCTGGAATTGTCAGTGGAAAAAGTTACTGAAGATCT TTTCACTTGCACTATGCTCTTTTATTGATTTGGCATCATTATCAAACACTG TTGGAGCCTTGTGAACTCTTGTTCAGAGTCTTCTGCCTCTCAAGGAATCAC ACTCC
>mGR15 aa MCAVLRSILTIIFILEFFIGNLGNG	>mGR15 nt AATAATAGATTTTTTAATATTCAGAATTTTTAAGTAATGTAGTATTGTTAG

FIALVQCMDLRKRRTF[redacted]ADHFLTA LAISRLALIWVLFLDSFLFIQSPLL MTRNTLRLIQTAWNISNHFSIWFAT SLSIFYLFKIAIFSNYLFFYLKRRV KRVVLVILLLSMILLFFNIFLEIKH IDVWIYGTKRNITNGLSSNSFSEFS RLILIPSLMFTLVPFGVSLIAFLLL IFSLMKHVRKMQYYTKGCKDVRTMA HTTALQTVVAFLLLYTTFFLSLVVE VSTLEMDESLMLLFAKVTIMIFPSI HSCIFILKHNKLRQDLLSVLKWLQY WCKREKTLDS	CAGCATAGCTTATAGGAAAAGTTCCAAGT[redacted]TTTGATTTTGTAATTCTGA TTCCCCCAAATCAAGTATCAAGTTTACCTGCACAGACAAGGGAAGAAGTGG CAAAATGTGCAAATGAGAGCAACTTTATTTGACTGTCAGTACGTTGAAATT CAGTGTTTCCTTAATCAGTTATGGATTGACATTTATGTGCACAGAACCTGG AAGAATTTCAGCCAAGCTGGAGGTAAAAATCCAAAATTCTGATGATAAAAC CAAAAGTAAATCACAGGTAAATCTTCTTTATTTTTCTTTTTTAATACTGTA TATGGACATTTTTTAATACAGCATATTTTTTTTTTGAAATTTAGAAAAAAA CCACTAAGAAATATTCACCAATGGAATAGACTTTAAAGTCACTTAGAGAAT GTGTGCTGTTCTACGTAGCATACTGACAATCATTTTCATTTTGGAGTTCTT CATTGGAAATCTGGGGAATGGATTCATAGCTCTGGTACAATGCATGGACTT ACGAAAGAGAAGAACGTTCCCTTCAGCAGATCATTTCCTCACTGCTCTGGC CATCTCCAGGCTTGCTCTGATATGGGTTTTATTTCTAGATTCATTTCTGTT TATACAATCCCCATTACTGATGACTAGAAATACATTAAGACTGATTCAGAC TGCCTGGAATATAAGCAATCATTTCAGTATATGGTTTGCTACCAGCCTCAG CATCTTTTATCTCTTCAAGATAGCCATTTTTTCTAACTATCTTTTCTTCTA CCTGAAGCGGAGAGTTAAAAGGGTGGTTTTGGTGATACTGCTGCTATCCAT GATCCTTTTGTTTTTTAATATATTTTTAGAAATCAAACATATTGATGTCTG GATCTATGGAACCAAAAGAAACATAACTAATGGTTTGAGTTCAAACAGTTT TTCAGAGTTTTCCAGGCTTATTTTAATTCCAAGTTTAATGTTCACATTAGT ACCCTTTGGTGTATCCTTGATAGCTTTCCTCCTCCTAATCTTTTCCCTTAT GAAACATGTAAGGAAGATGCAGTACTACACCAAAGGATGCAAAGATGTCAG AACCATGGCCCACACCACAGCCCTGCAGACTGTGGTTGCCTTCCTCCTATT ATATACTACTTTCTTTCTGTCTCTAGTTGTGGAAGTTTCAACACTTGAAAT GGATGAAAGTCTGATGCTTCTGTTTGCAAAAGTTACTATAATGATTTTTCC TTCCATCCACTCCTGTATTTTCATTTTGAAACATAATAAGTTGAGACAGGA CTTGCTTTCAGTACTGAAGTGGCTACAGTATTGGTGCAAGCGTGAGAAAAC CTTGGATTCATAGACCATTGTATGCATCACCTTGAATATTCTAGAGGGGTG TAGGTTCATATGAAAGTATTGAATTTTTAAATTTGAGCCTTTTGTATATTT TCT
>mGR16 aa MNGVLQVTFIVILSVEFIIGIFGNG FIAVVNIKDLVKGRKISSVDQILTA LAISRIALLWLILVSWWIFVLYPGQ WMTDRRVSIMHSIWTTFNQSSLWFA TSLSIFYFFKIANFSNPIFLYLKVR LKKVMIGTLIMSLILFCLNIIIMNA PENILITEYNVSMSYSLILNNTQLS MLFPFANTMFGFIPFAVSLVTFVLL VFSLWKHQRKMQHSAHGCRDASTKA HIRALQTLIASLLLYSIFFLSHVMK VWSALLLERTLLLLITQVARTAFPS VHSWVLILGNAKMRKASLYVFLWLR CRHKE	>mGR16 nt TTTATGATGGAAAGAATAAAACCATTAGCAAGGCTTAATGGCTTGTTTGGT ATTAGACCTGTACATTGTTTATGGAACATGATATGGAGCTTTGTTTATTGA ATATGCACAATATTTTAGAAGCATGTTTCAAAGAATCTTAAGTAATTACAA TAGAAATTGAAGCATCCAAGTGAAGATGAATGGTGTCCTACAGGTTACATT TATAGTCATTTTGAGTGTGGAATTTATAATTGGCATCTTTGGCAATGGATT CATAGCGGTGGTGAACATAAAGGACTTGGTCAAGGGAAGGAAGATCTCTTC AGTGGATCAGATCCTCACTGCTCTGGCCATCTCCAGAATTGCACTGCTGTG GTTAATATTAGTAAGTTGGTGGATATTTGTGCTTTACCCAGGACAATGGAT GACTGATAGAAGAGTTAGCATAATGCACAGTATATGGACAACATTCAACCA GAGTAGTCTCTGGTTTGCTACAAGTCTCAGCATCTTTTATTTTTTCAAGAT AGCAAATTTTTCCAACCCTATTTTTCTTTATTTAAAGGTCAGACTTAAAAA AGTCATGATAGGGACATTGATAATGTCTTTGATTCTCTTTTGTTTAAATAT TATCATTATGAATGCACCTGAGAACATTTTAATCACTGAATATAATGTATC TATGTCTTACAGCTTGATTTTGAATAACACACAGCTTTCTATGCTGTTTCC ATTTGCCAACACCATGTTTGGGTTCATACCTTTTGCTGTGTCACTGGTCAC TTTTGTCCTTCTTGTTTTCTCCCTGTGGAAACATCAGAGAAAGATGCAACA CAGTGCCCATGGATGCAGAGATGCCAGCACTAAGGCCCACATCAGAGCCTT GCAGACATTGATTGCCTCCCTCCTCCTGTATTCCATTTTCTTCCTGTCTCA TGTTATGAAGGTTTGGAGTGCTCTGCTTCTGGAGAGGACACTCCTGCTTTT GATCACACAGGTTGCAAGAACAGCTTTTCCGTCAGTGCACTCCTGGGTCCT GATTCTGGGCAATGCTAAGATGAGAAAGGCTTCTCTCTATGTATTCCTGTG GCTGAGGTGCAGGCACAAAGAATGAAACCCTACAGTGTACAGACCTGGGGT ATATTTATGTGGATGATCTTACATATCTTAGAGGAAAATGGATTAAAAGAA ATTCTCATATTTATAAATTTTTAGGTCTGAATTACATAAAAATGTATATAA TATTTTCAAAGTACAAGATAGTAGTTTATAACTTACATGATAAATACTGTC TATGCATCTTCTAGTCTTTGTAGAATATGTAAAAACATGTT
>mGR17 aa MKHFWKILSVISQSTLSVILIVELV IGIIGNGFMVLVHCMDWVKKKKMSL VNQILTALSISRIFQLCLLFISLVI NFSYTDLTTSSRMIQVMYNAWILAN HFSIWIATCLTVLYFLKIANFSNSF	>mGR17 nt GAATTCTGGTCTGGCACCCCTGAGCTGTGTGAGTAGACACATTATCATGGA AAGAGATTCAGAATCTGTCACTGTCAAAACTGCATGTTTGCTCCTCTGTTA GTGTGTTGGGGAAAGTTAAGAAAAATACATTTTATGAGAATCAACTCAGAG GTTGTCAGAAATTGTCGAAACAGCATTTTAAAAATTTACATCTCAACTGGA TATATGAGCAAGTCTTTATAACTGATATATAAAATGAAGCACTTTTGGAAG ATATTATCTGTTATCTCCCAGAGCACACTTTCAGTCATTTTAATCGTGGAA

FLYLKWRVEKVVSVTL[redacted]LLLLIL NILLTNLETDMWTNEYQRNISCSFS SHYYAKCHRQVLRLHIIFLSVPVVL SLSTFLLLIFSLWTHHKRMQQHVQG GRDARTTAHFKALQTVIAFFLLYSI FILSVLIQIWKYELLKKNLFVVFCE VVYIAFPTFHSYILIVGDMKLRQAC LPLCIIAAEIQTTLCRNFRSLKYFR LCCIF	TTAGTAATTGGAATTATAGGAAATGGGTT[redacted]GGTCCTGGTCCACTGTATG GACTGGGTTAAGAAAAAGAAAATGTCCCTA[redacted]TAATCAAATTCTTACTGCT TTGTCAATCTCCAGAATTTTTCAGCTCTGTTTATTGTTTATAAGTTTAGTA ATCAACTTTTCATATACAGATTTAACTACAAGTTCAAGGATGATACAAGTC ATGTACAATGCTTGGATTTTAGCCAACCATTTCAGCATCTGGATTGCTACA TGCCTCACTGTCCTTTATTTTCTAAAGATAGCCAATTTTTCTAACTCTTTT TTTCTTTATCTAAAGTGGAGAGTTGAAAAAGTAGTTTCAGTTACACTGTTG GTGTCATTGCTCCTCCTGATTTTAAATATTTTACTAACTAACTTGGAAACC GACATGTGGACAAATGAATATCAAAGAAACATATCATGCAGCTTCAGTTCT CATTACTATGCAAAGTGTCACAGGCAGGTGTTAAGGCTTCACATTATTTTC CTGTCTGTCCCCGTTGTTTTGTCCCTGTCAACTTTTCTCCTGCTCATCTTC TCCCTGTGGACACATCACAAGAGGATGCAGCAGCATGTTCAGGGAGGCAGA GATGCCAGAACCACGGCCCACTTCAAAGCCCTACAAACTGTGATTGCATTT TTCCTACTATATTCCATTTTTATTCTGTCTGTCTTAATACAAATTTGGAAA TATGAATTACTGAAGAAAAATCTTTTCGTTGTATTTTGTGAGGTTGTATAT ATAGCTTTTCCGACATTCCATTCATATATTCTGATTGTAGGAGACATGAAG CTGAGACAGGCCTGCCTGCCTCTCTGTATTATCGCAGCTGAAATTCAGACT ACACTATGTAGAAATTTTAGATCACTAAAGTACTTTAGATTATGTTGTATA TTCTAGACAAAAATTAACTGATACAAATGTCTTTTGTATTTTTCATTTTAA ATATCCTTTAATTTTGACTGCATGAAATTGATTTCTGCTTGCAATTATCAC TGATTAAAACTATTAATAATTTAACTAG
>mGR18 aa MVPTQVTIFSIIMYVLESLVIIVQS CTTVAVLFREWMHFQRLSPVETILI SLGISHFCLQWTSMLYNFGTYSRPV LLFWKVSVVWEFMNILTFWLTSWLA VLYCVKVSSFTHPIFLWLRMKILKL VLWLILGALIASCLSIIPSVVKYHI QMELVTLDNLPKNNSLILRLQQFEW YFSNPLKMIGFGIPFFVFLASIILL TVSLVQHWVQMKHYSSSNSSLKAQF TVLKSLATFFTFFTSYFLTIVISFI GTVFDKKSWFWVCEAVIYGLVCIHF TSLMMSNPALKKALKLQFWSPEPS	>mGR18 nt GCGTGCTTCACAGAGCAGTATACTACAAAGCAAATGTCATTGCTGCCATTG TATATTTCTCTAAAGACATTTCACATTTTATCTCCCTGTCCCATTGTGTGC AGAGCCCACACTTCAATCAATCAATTCCTTAATTATAAGCTATTGTTTCAT TATTTCATTTCCTACGTTTTTTTGCATTTTTACTAAAACTCCAAAGCAGAC ATTTTCTAATTATAATCCTACATGTAGTTAGAATTTTAAAAATTATATACT ATTTTCTTTGCACCACTGAGTTCAGTAGGTTTTGAAGGTTTATGCTTAACA ATTGAACATTTCATGTTAGATTATTCCTGCCTTCCTAATCTTGAATAATTA AATGTCCATCCAGGCTTAGAATTCACAGAGTCAACAGCTTTCACCTTGATT CTCTCACTATCTATCAATGACTAGAATCTGTCTGTCACTTTTGAAACCGCT AATTAAATAGTTGGTGCTTATTTAAAGGGTGCCCCATGCCAAGAGAAAATG TATTTCTTCTCTAGATGCCTTCGTCCTTTACAAGTTACATGCTTTACTGAT GGTGAATTGGTTTTCTTCCAGTTCATCTGGGTTAAGTGACCTAAGAACCTA GCCATGGAAGGAGAAACAGAAGCAAATATTAACGATACAAGAACAAGTTCC AGAACATTGGAAAGTACTTAGTAAAGGCATTGGAATTAGCAAAAGAATAGT AGCGAAGCAAAAAATACTTCATCTCCATTGGGAGGTCAAGAAAGACTATGC AGTGTTTTTGATGCAACTTGTCATCTCTGAGTTAGACGATTCAGCACACAC TTTTGAGATTGAACTTCAACAGGTGGAGCCAGCAGACCTGAGCTTTAGGAA TGATGGTGGAATTTCCAAGCAAAGACTTCCGTTACCTTTTTGATGTCCCCT AACAATTCGGTTGCAATGCTCACACCGCCCAACTGTTGAAATGCTTGGGAA AAGGGATTCTGAGACTGGCATTAGTATGTCATTTGACAGAATGGAAACATT GCCCAGGGCATTAATGCACAGTAAAGGATTCACCTTTTCTAAGTGCTCAAA TTTTAAATTTGnATATTTTTAGAAGACATTATTTAAAAGAAAGGTGGAGAG GATATCCAAACAGCACCTTGAGCAGATAAAGAGGTGAAGAAGAAAAAACAA CATGCGTACATGATGGATTTCTCTTTATGAAAATGATCAAATGATCTTAGG ATCAAGAATCCACACCTGAATGAGATTTGCTTGTATCCCTGTGTGAATTTG ACCTAACAAGCAAAGCACAGACAAATGCTGTAGATAGGGAAATGTCTATGT CAAATGTGTGTAAGGAGGATTTGCATCCACAAAGAAGTGCCCTCTTATACT GAGAGTGCTAAGAACACATGTCCGTTTCATATTCGGAAAGTGGTATAGAGC TGTTGAGTCTTTGGCTAGGAAGAGACTTCAGAGTGGAAGCATGGTGCCAAC GCAAGTCACCATCTTCTCCATCATCATGTATGTGCTTGAGTCCTTAGTAAT AATTGTGCAAAGTTGCACAACGGTTGCAGTGCTATTCAGAGAGTGGATGCA CTTTCAAAGACTGTCACCGGTGGAGACGATTCTCATCAGCCTGGGCATCTC ACATTTCTGTCTACAGTGGACATCAATGCTATACAACTTTGGTACTTATTC TAGGCCTGTCCTTTTATTTTGGAAGGTATCAGTCGTCTGGGAGTTCATGAA CATTTTGACATTCTGGTTAACCAGTTGGCTTGCTGTCCTCTACTGTGTCAA GGTCTCTTCCTTCACTCACCCCATCTTCCTCTGGCTGAGGATGAAAATCTT GAAACTGGTTCTCTGGTTGATACTGGGTGCTCTGATAGCTTCTTGTTTGTC AATCATCCCTTCTGTTGTTAAATATCACATCCAGATGGAATTAGTCACCCT AGATAATTTACCCAAGAACAATTCTTTGATTCTAAGACTACAACAGTTTGA ATGGTATTTTTCTAATCCTTTAAAAATGATTGGCTTTGGTATTCCTTTCTT CGTGTTCCTGGCTTCTATCATCTTACTCACAGTCTCATTGGTCCAACACTG

	GGTGCAGATGAAACACTACAGCAGCAGCAA[redacted]CCAGCCTGAAAGCTCAGTT CACTGTTCTGAAGTCTCTTGCTACCTTCTT[redacted]CTTCTTCACATCCTATTT TCTGACTATAGTCATCTCCTTTATTGGCACTGTGTTTGATAAGAAATCTTG GTTCTGGGTCTGCGAAGCTGTCATCTATGGTTTAGTCTGTATTCACTTCAC TTCACTGATGATGAGCAACCCTGCATTGAAAAAGGCACTGAAGCTGCAGTT CTGGAGCCCAGAGCCTTCCTGAGGCAGGAAACACAGTTAAGCCTCTAGGGT AAGGAGACTTTGCATTGGCACAGTCCCTATAGTGTAATGCAAACTTGAACA CAAACTTCATCCCTTTTCACATCCACAAATGGCTGCATCTATACATCATCA CCAGTCTTCCCTGTATTCTGACCCATTCTCTTCCTGTCCTATCCATAGTCC CCAGGTTGGTTTTGATTTTTCTCATGATCACACCAACTCTGCTTAGCTTTT GCCACCACTGTAATAGTAAACATGGGGTGTTCTATATATTACAGTCAAAAT CATTCTCACATTGTTGATTGCCTCACAAATTCATATAAATCCCCCTTCCTG TCAGGAATTTATTGTCTGCTCACTTAATGCTCACCATATATTAAAGCCATT AATTCCCCCTTCCTACCTTGAGTTTAAGAAGGAAAATGTCTTACCATTGCC CACAACCTATTCTGCTGCTTCTAGACTTTTATGCAAGTGATTTATACACAC ACACACACACACACACACACATACAAACAAC
>mGR19 aa MMEGHMLFFLLVVVVQFLTGVLANG LIVVVNAIDLIMWKKMAPLDLLLFC LATSRIILQLCILFAQLGLSCLVRH TLFADNVTFVYIINELSLWFATWLG VFYCAKIATIPHPLFLWLKMRISRL VPWLILASVVYVTVTTFIHSRETSE LPKQIFISFFSKNTTRVRPAHATLL SVFVFGLTLPFLIFTVAVLLLLSSL WNHSRQMRTMVGTREPSRHALVSAM LSILSFLILYLSHDMVAVLICTQGL HFGSRTFAFCLLVIGMYPSLHSIVL ILGNPKLKRNAKTFIVHCKCCHCAR AWVTSRNPRLSDLPVPATHHSANKT SCSEACIMPS	>mGR19 nt CTGCAGCCTAGAGAACTAATGCATAGGAAACTTATATTCCCACCTCCGTGA CGTCACTCTGACAGAAGTGAACTTATATTCCCACCTCCGTGACGTCACTCT GACAGAAGTGACTTGTTTTTGTATGATGCTCCAGGATGCCTCATTAGCATT GAGGACAATCATAATTAAGTAAGGCAAGGCATGAAGGTGGTCCTCACTAGG TACCTGGAGGCTTCTGGTTGCATGATTTACTTGTGATGACTCTGACACTTA AGAAGACCTGAAAAATGCAAAAGCTGTCATAAGGCACAGTTCGTTTCTATG GTATCTCTTCCTTATTTGACTGACATTGAGTTGAGAAGGCAGCACTATAAA CAAATGGGCCCCACCTTCCTCTTCCATTGTCTTTGGGTTGGCATCATCTCC AAAGGAACCTTGGTCTAGTTGAAAGAAGCCAGAAATCATACATGGCTGAGA CTGTGCATAACTCTATGTATCATTTAAAGAAGTCATTGGTTCTTCTTATTT TAAAATGATGGAAGGTCATATGCTCTTCTTCCTTCTGGTCGTGGTAGTGCA GTTTTTAACTGGGGTCTTGGCAAATGGCCTCATTGTGGTTGTCAATGCCAT CGACTTGATCATGTGGAAGAAAATGGCCCCACTGGATCTGCTTCTTTTTTG CCTGGCGACTTCTCGGATCATTCTTCAATTGTGTATATTGTTTGCACAGCT GGGTCTATCCTGTTTGGTGAGACACACGTTATTTGCTGACAATGTTACCTT TGTCTACATTATAAACGAACTGAGTCTCTGGTTTGCCACATGGCTTGGTGT TTTCTACTGTGCCAAGATTGCTACCATCCCTCACCCACTCTTTCTGTGGCT GAAGATGAGGATATCCAGGTTGGTGCCATGGCTGATCCTGGCATCTGTGGT CTATGTAACTGTTACTACTTTCATCCATAGCAGAGAGACTTCAGAACTTCC TAAGCAAATCTTTATAAGCTTTTTTTCTAAAAATACAACTCGGGTCAGACC AGCGCATGCCACACTACTCTCAGTCTTTGTCTTTGGGCTCACACTACCATT TCTCATCTTCACTGTTGCTGTTCTGCTCTTGTTGTCCTCCCTGTGGAACCA CAGCCGGCAGATGAGGACTATGGTGGGAACTAGGGAACCTAGCAGACATGC CCTCGTCAGTGCGATGCTCTCCATTCTGTCATTCCTCATCCTCTATCTCTC CCATGACATGGTAGCTGTTCTGATCTGTACCCAAGGCCTCCACTTTGGAAG CAGAACCTTTGCATTCTGCTTATTGGTTATTGGTATGTACCCCTCCTTACA CTCGATTGTCTTAATTTTAGGAAACCCTAAGCTGAAACGAAATGCAAAAAC GTTCATTGTCCATTGTAAGTGTTGTCATTGTGCAAGAGCTTGGGTCACCTC AAGGAACCCAAGACTCAGCGACTTGCCAGTGCCTGCTACTCATCACTCAGC CAACAAGACATCCTGCTCAGAAGCCTGTATAATGCCATCTTAATTGTCCAA CCTGAGGCTTAATCATTTCAAAGGGTAAATTGATGATCAAAGCCCAACACA TGATATGACATCAAGGTCCATATCCCAGTAGTCATGTGGAAATACCACCTT GCAAAATGATGTCATTGAGAAACCAGGGCAAATGGAGTCTAGGTCTTTCAG TATGATTTGCTGCAG
>mGR20 aa MNLVEWIVTIIMMTEFLLGNCANVF ITIVNFIDCVKRRKISSADRIITAI AIFRIGLLWAMLTNWHSHVFTPDTD NLQMRVFGGITWAITNHFTTWLGTI LSMFYLFKIANFSNSLFLHLKRKLD NVLLVIFLGSSLFLVAYLGMVNIKK IAWMSIHEGNVTTKSKLKHVTSITN MLLFSLINIVPFGISLNCVLLLIYS LSKHLKNMKFYGKGCQDQSTMVHIK ALQTVVSFLLLYATYSSCVIISGWS	>mGR20 nt CTAGATGGGCTGTTTCATATAATGACTGGAACTCCCTACATGCTCCACGTC TTGAGTTCTAAAATTTCACTAACAAATTTTTGACTGCCATAAATAATGAAG GTTTAAAGAAAGAACAACATTTGAAGCAATGGACCAGAATTCCTCTTTATT TGACTCTTAGCAAATTGGAATGCAGCATCCTTTCAAGAGCAGCACTGAAAT ATACCAGTCAATGGCAGAGAGTAAAAAAGTATGCAATTGGAGACATTATGG TAATATAAATTTCCATTAAAAATGAGACTGCATTCACCTATTACAACACAT TGCTATTCTGCTCAACACAGAGTTAAAAAGAAACAAGAACTCTTGTATACA TTCAGTTAGTCACAAGTATAATTATGTTCACATATTTTAAAAAAATGAATC ATGATCTGTGAATTGAGCCTGGCTTTTTTTGTCTCTCTCTTTTTATTCTTT TCCTTTAGACAGACACAATGAATTTGGTAGAATGGATTGTTACCATCATAA TGATGACAGAATTTCTCTTAGGAAACTGTGCCAATGTCTTCATAACCATAG

LQNAPVFLFCVTIGSF[redacted]GHSCIL IWGNQKLKQVFLLLLRQMRC	TGAACTTCATCGACTGTGTGAAGAGAAGA[redacted]ATCTCCTCAGCTGATCGAA TTATAACTGCTATTGCCATCTTCAGAATTGG[redacted]TTGTTGTGGGCAATGTTAA CGAACTGGCATTCACATGTGTTTACTCCAGACACAGACAATTTACAAATGA GAGTTTTCGGTGGAATTACCTGGGCTATAACCAACCATTTTACCACTTGGC TGGGGACCATACTGAGCATGTTTTATTTATTCAAGATAGCCAATTTTTCCA ACAGTCTATTTCTTCATCTAAAAAGAAAACTTGACAATGTTCTACTTGTGA TTTTCCTGGGATCGTCTCTGTTTTTGGTTGCATATCTTGGGATGGTGAACA TCAAGAAGATTGCTTGGATGAGTATTCATGAAGGAAATGTGACCACAAAGA GCAAACTGAAGCATGTAACAAGCATCACAAATATGCTTCTCTTCAGCCTGA TAAACATTGTACCATTTGGTATATCACTGAACTGTGTTCTGCTCTTAATCT ATTCCCTGAGTAAACATCTCAAGAATATGAAATTCTATGGCAAAGGATGTC AAGATCAGAGCACCATGGTCCACATAAAGGCCTTGCAAACTGTGGTCTCTT TTCTCTTGTTATATGCCACATACTCTTCCTGTGTCATTATATCAGGTTGGA GTTTGCAAAATGCACCAGTCTTCCTGTTTTGTGTGACAATTGGATCCTTCT ACCCAGCAGGTCATTCTTGTATCTTGATTTGGGGAAACCAGAAACTTAAAC AGGTCTTTCTGTTGTTGCTGAGGCAGATGAGATGCTGACTGAAAAAATGAA AGTCCCCCTGTCTCTAG
>mGR21 aa MGSNVYGILTMVMIAEFVFGNMSNG FIVLINCIDWVRKGTLSSIGWILLF LAISRMVLIWEMLITWIKYMKYSFS FVTGTELRGIMFTWVISNHFSLWLA TILSIFYLLKIASFSKPVFLYLKWR EKKVLLIVLLGNLIFLMLNILQINK HIEHWMYQYERNITWSSRVSDFAGF SNLVLLEMIVFSVTPFTVALVSFIL LIFSLWKHLQKMHLNSRGERDPSTK AHVNALRIMVSFLLLYATYFISFFL SLIPMAHKTRLGLMFSITVGLFYPS SHSFILILGHSNLRQASLWVMTYLK CGQKH	>mGR21 nt CTCTTTTGAAGACAATAGTTGTTCTACTAGCTATTGATAGCATGTTTACAT TTGTCATTTTCAAGTATGTTCAGAAACAAAGCTACATATTGTGGGGAGTAT ATAAAATATGAAAGCATGCCATTCCCAGGCATCCAAGGATCCCTGTGTATT AAAAGGCAACAAAGCAGAACCAAATGTTCTGTTTTGGACATGAGCTTCTTC CAATTCAACTGCTGAAAAATTTGGATAACTACATATAAAACTAAGAACACA GAGTGTCACAGAGCAGTCTCTGCTCTCCAATTCACCAGGATTAATATTGAC AGACCCAAAAGATGTCATTTAGGTAAATTTTGGATGAATCATATTGTTGTC ACCTTTGTGCTCTAGAACATAAGCTGATAGAATCAAATTTTCTTTAGCAGA GACAATGCAAATTGATATAACAGTGAAAGAGAATATATCTTTATTTGCATG TTAGCAAATGACAGCTGGATGCACTTCATGATTTTCTGCAATCTAGTTCAG TCTTTAGAAGGATATATATATATATATATATATATATATATATATATATAT ATATATATATATAAACCTTAGTCTTGAAAGATATCAGAAAGAAGGATTTCA CAAGAATGTACAGAGCCATTAGCAAAATTTTAATATACTCATCGACATTAG GTCAGTCACTACATAAGAAGGACTTGAATGAAAGCTTATCTTAGTTTTTGA GACTACAGGGACATTTCACCTTGCCAAATGAGAAGCAGTGAGTCTTCTTTG TCTGGACATGGGAAGCAATGTGTATGGTATCTTAACTATGGTTATGATTGC AGAGTTTGTATTTGGAAATATGAGCAATGGATTCATAGTGCTGATAAACTG CATTGATTGGGTCAGGAAAGGAACTCTTTCTTCCATTGGTTGGATCCTGCT TTTCTTGGCCATTTCAAGAATGGTGTTGATATGGGAAATGTTAATAACATG GATAAAATATATGAAGTATTCATTTTCATTTGTGACTGGAACAGAATTACG GGGTATCATGTTTACCTGGGTAATTTCCAATCACTTCAGTCTCTGGCTTGC CACTATTCTCAGCATCTTTTATTTGCTCAAAATAGCCAGTTTCTCCAAACC GGTTTTTCTCTATTTGAAGTGGAGAGAGAAGAAAGTGCTTCTGATTGTCCT TCTGGGAAATTTGATCTTCTTGATGCTCAACATATTACAAATAAACAAACA TATAGAACACTGGATGTATCAATATGAGAGAAATATAACTTGGAGTTCTAG AGTGAGTGACTTTGCAGGGTTTTCAAATCTGGTCTTATTGGAGATGATTGT GTTCTCTGTAACACCATTCACAGTGGCCCTGGTCTCCTTCATCCTGTTAAT CTTCTCCTTGTGGAAACATCTACAGAAAATGCATCTCAATTCTAGAGGGGA ACGAGACCCCAGCACTAAAGCCCATGTGAATGCCTTGAGAATTATGGTCTC CTTCCTCTTACTCTATGCCACTTACTTCATATCTTTTTTTCTATCATTGAT TCCCATGGCACATAAAACACGACTGGGTCTTATGTTTAGCATAACTGTTGG GCTTTTCTACCCTTCAAGCCACTCATTTATCTTAATTTTGGGACATTCTAA TTTAAGGCAAGCCAGTCTTTGGGTGATGACATATCTTAAATGTGGGCAAAA GCATTAGAATTTCACTATTCCATAAGGCAGCCAAACCACGTGCTACTAGGT ATATGATACTACTCAGTGGTAAAGCCCTAGGCAAACATTAACCTTAGAAAA TATATAATTTTGTGACTCTTCTGTATTTGATAAATCACTCACATATTTAGA AGAATGCTACAGTAGTGTGATCTTGTACATGATTGTAACAATTCAATTTTA TTAATATAGTTCAGGCATGATAACATACCCCTGATAACTGAAAAGTAAGTA GGATGCTACATATATATTTAGATCTAGACTTAGGGGCAAAGAGAGACCCAG CTGATAGCTGTGCAATAAAGATTTTAATTTTCATCCTGTTGTGAGTTATCT GAAATCTATGTCACTGAAGGCATAAGCAAGATTTTCACACACTGAAACAAT CTCTTATGCTTTCTTATATTGTTTTAAAAGTAAATTAGAAAATTTAAATAA ACTTAATGGCAATTGAAATTACAAAAGCTAAACACATGTGGTTATTAGAAA TTAGACTGTATGTAGGTCCTAGGGGATGGCTTAGTAAAGTGCTTTGTTGCA AGCTTCAGGATATGATTCTAAATCCCTAGATTCAATTAAAAACCTGGCATA

	AATAGCCAATGTAAAATTTGTCTGTAAA[redacted]TAACCAGTGCTAAGAGTACC AAGACAACAAAATGTTTACTTTTAAAACCATTTATTGATATTCTTTTAAAA ATAGGTATGTATTTTACTATTTAAATAAGATTTTGTCAAAAGCTAGTCTTG ACACCTTAGGTAAACATAGGAAGGCAACAAGTTTGAAGTCAGCTACTGGGG ACAGTGCTGCTAGCAGCTGACAGAGGCCACTGCTGACTACAGCAGATCATT TACAGGTTCAGCACTAG
>mGR22 aa MSSLLEIFFVIISVVEFIIGTLGNG FIVLINSTSWFKNQKISVIDFILTW LAISRMCVLWTTIAGASLRKFYKTL SYSKNFKFCFDIIWTGSNYLCIACT TCISVFYLFKIANFSNSIFFWIKQR IHAVLLAIVLGTLMYFILFLIFMKM IANNFIYKWTKLEQNTTFPVLDTLS GFLVYHSLYNGILIFFFIVSLTSFL LLIFSLWSHLRRMKLQGIHTKDIST EAHIKAMKTMMSFLLFFIIYYISNI MLIVASSILDNVVAQIFSYNLIFLY LSVHPFLLVLWNSKLKWTFQHVLRK LVCHCGGYS	>mGR22 nt AAATGAATAATTTCATGCAAAGGATACCATTAGAATATGATCACTATTTAA ATTTTAGCAAATACATATTCAAATACCAGCACAATGTTTCAAATTTAAAAT ATAAACATTATAAAACCCAGCAGAGAACAAAATGATAGCCTTGATAATTGT TGGTTTGCTCAAGAAAAATGGGTGTATACTTTAACATTTAATTGGGAACTC AGTTGAGAGCATACATTTAGGGTTTTACAGAGGTATTCATTGCCCATTTAA GATTTGGATTCACACATCTACATCAATGTGGCTGTAATCCATTTTCCCATG ATGAAATAAGGTAGAGACTGCCTATTAAACGACATGTCGAGCCTACTGGAG ATTTTCTTTGTGATCATTTCGGTTGTAGAATTCATAATAGGAACTTTGGGA AATGGATTTATTGTCCTGATAAACAGTACTTCTTGGTTCAAGAATCAGAAA ATCTCTGTAATTGATTTCATTCTTACTTGGTTGGCCATCTCCAGAATGTGT GTTCTATGGACAACAATTGCTGGTGCCTCTCTCAGGAAATTCTACAAGACG TTAAGTTACTCTAAGAATTTCAAATTTTGTTTTGACATTATCTGGACAGGA TCCAACTATTTATGCATAGCCTGTACAACGTGCATCAGTGTCTTCTACTTG TTCAAGATTGCCAACTTTTCTAATTCCATTTTCTTCTGGATTAAACAGAGA ATTCATGCAGTACTTCTGGCTATTGTCCTAGGCACACTCATGTATTTCATT TTATTTCTCATTTTTATGAAAATGATAGCTAATAATTTTATCTACAAATGG ACAAAATTGGAACAAAACACAACATTCCCTGTTTTAGATACTCTAAGTGGT TTCTTAGTCTACCATAGCCTCTACAATGGGATTCTCATTTTCTTTTTTATA GTGTCTCTGACCTCATTTCTTCTTTTAATCTTCTCTTTATGGAGCCACCTT AGGAGGATGAAACTACAGGGCATACATACCAAAGACATAAGCACAGAAGCA CACATAAAAGCTATGAAAACTATGATGTCATTCCTTTTGTTCTTCATCATA TATTATATTAGCAACATTATGCTTATTGTGGCAAGCTCCATTCTTGACAAT GTGGTTGCACAAATTTTCTCTTATAACCTAATATTTCTGTATTTATCTGTT CATCCTTTCTTCTGGTTTTATGGAACAGCAAATTGAAATGGACATTCCAG CATGTATTGAGAAAGCTGGTGTGTCATTGTGGAGGTTATTCTTGATTTCAG TAAATACACTCAATATAACTGATGGATTTCTAAGGTAAGAAAAATGGAACA AGGAATAAAGAGGAGAAATATATTCCTTTTCAGATCATCTGCTCTGTCATT CTGTCCTTAGCATGCTATTAAGAATTGTTGACTAAATCCAGTCATTTTTAA CATGAGGAAAGGATGTTTCAATCCAACTTAGAGAGGGTACAAAATAGTCCT AGGAGGCAG
>mGR23 aa MFSQKINYSHLFTFSITLYVEIVTG ILGHGFIALVNIMDWVKRRRISSVD QILTALALTRFIYVLSMLICILLFM LCPHLPRRSEMLSAMGIFWVVNSHF SIWLTTCLGVFYFLKIANFSNSFFL YLKWRVKKVILIIILASLIFLTLHI LSLGIYDQFSIAAYVGNMSYSLTDL TQFSSTFLFSNSSNVFLITNSSHVF LPINSLFMLIPFTVSLVAFLMLIFS LWKHHKKMQVNAKQPRDVSTMAHIK ALQTVFSFLLLYAIYLLFLIIGILN LGLMEKIVILIFDHISGAVFPISHS FVLILGNSKLRQASLSVLPCLRCQS KDMDTMGL	>mGR23 nt AATTTTCAGCAACCAATATGTAGACTGCTTAAATGCATCAGAAACATTATA AATTGAAGCATGTTTTCACAGAAAATAAACTACAGCCATTTGTTTACTTTT TCAATCACCTTGTATGTGGAAATAGTAACGGGAATCTTAGGACATGGATTC ATAGCATTAGTGAACATCATGGACTGGGTCAAAAGAAGAAGGATCTCTTCA GTGGATCAGATTCTCACTGCTTTGGCCCTTACCAGATTCATTTATGTCTTG TCTATGCTGATTTGCATATTGTTATTCATGCTGTGCCCACATTTGCCTAGG AGATCAGAAATGCTTTCAGCAATGGGTATTTTCTGGGTAGTCAACAGCCAT TTTAGCATCTGGCTTACTACATGCCTCGGTGTCTTTTATTTTCTCAAGATA GCCAATTTTTCTAACTCTTTTTTTCTTTATCTAAAGTGGAGAGTTAAAAAA GTGATTTTAATAATAATCCTGGCATCACTGATTTTCTTGACTTTACACATT TTATCTTTAGGGATATATGATCAGTTCTCAATTGCTGCTTATGTAGGAAAT ATGTCTTATAGTTTGACAGATTTAACACAATTTTCCAGTACTTTCTTATTC TCCAACTCATCCAATGTTTTCTTAATCACCAACTCATCCCATGTTTTCTTA CCCATCAACTCCCTGTTCATGCTCATACCCTTCACAGTGTCCCTGGTAGCC TTTCTCATGCTCATCTTCTCACTGTGGAAGCATCACAAAAAGATGCAGGTC AATGCCAAACAACCTAGAGATGTCAGTACTATGGCCCACATTAAAGCCTTG CAAACTGTGTTCTCCTTCCTGCTGCTGTATGCCATATACTTACTTTTCCTT ATCATAGGAATTTTGAACCTTGGATTGATGGAGAAAATAGTGATACTGATA TTTGACCACATTTCTGGAGCAGTTTTTCCTATAAGCCACTCATTTGTACTG ATTCTGGGAAACAGTAAGCTGAGACAAGCCAGTCTTTCTGTGTTGCCTTGT CTAAGGTGCCAGTCCAAAGATATGGACACCATGGGTCTCTAGTAAATTCCA GAGTACATTTTGTAAAAATCTTGAGGATGATCAGTTCATAGAAAAAAGTTA CCTTATGGGGGAAAATAAAAAGTGGGGCTTCAATCCTGGGAGTAATAATAC ACAGGAGGGTAGGACAGCATGAAGGAGACTAGCACTATATAAGTGGTCTCA TACAGGATATGGGAAAGGAAAGATTTATGCAATAAAGAGGGAGATCATATT

	GGAGGATGAGGAGGCATTACATATGTAAA[redacted]ACTATAAGAATGGAATCAT GCTAATCTAAAAAAATCTGTAATGCATTTCATTCAGACTATATACATATAT GCCTATATATGGATATATGGGGATATATATTCTATACATATTTTAAAAGAA CCTTTCTTATATAG
>mGR24 aa MVPVLHSLSTIILIAEFVWGNLSNG LIVLKNCIDWINKKELSTVDQILIV LAISRISLIWETLIIWVKDQLISSI TIEELKIIVFSFILSSHFSLWLATA LSIFYLFRIPNCYWQIFLYLKWRIK QLIVHMLLGSLVFLVANMIQITITL EERFYQYGGNTSVNSMETEFSILIE LMLFNMTMFSIIPFSLALISFLLLI FSLWKHLQKMPLNSRGDRDPSATAH RNALRILVSFLLLYTIYFLSLLISW VAQKNQSELVHIICMITSLVYPSFH SYILILGNYKLKQTSLWVMRQLGCR MKRQNTPTT	>mGR24 nt CAAAGAGGAGAAATATTTAGCTACACAGTGTACCACATACAAGCCGTTCAA TCAGTATAAGGGGAGCAGTCATATAGAATTTGGGCTTTCTTTCTTTTAATA TGGTACCTGTTCTGCACAGTCTCTCCACCATCATACTAATTGCAGAGTTTG TTTGGGGAAATTTGAGCAATGGTTTGATAGTGTTGAAGAACTGCATTGACT GGATCAATAAAAAAGAGCTCTCCACAGTTGATCAAATACTCATTGTCTTGG CAATTTCAAGAATTAGTCTCATCTGGGAAACACTAATTATATGGGTTAAAG ATCAACTAATTTCATCTATTACTATTGAAGAATTAAAAATAATTGTGTTCA GCTTTATACTATCTAGCCACTTCAGTCTCTGGCTTGCTACAGCTCTCAGCA TCTTCTATTTATTCAGAATACCTAATTGCTACTGGCAGATCTTTCTCTACT TGAAATGGAGAATAAAGCAACTGATTGTCCACATGCTTCTGGGAAGCTTGG TGTTCTTGGTTGCAAATATGATACAGATAACCATCACTCTTGAAGAGAGGT TCTATCAATATGGAGGAAATACAAGTGTAAATTCCATGGAGACTGAGTTCT CAATTTTGATAGAGCTGATGTTATTTAACATGACTATGTTCTCCATTATAC CATTTTCATTGGCCTTAATTTCTTTTCTTCTGCTAATCTTCTCTTTATGGA AACATCTCCAGAAGATGCCACTCAATTCTAGAGGAGATAGAGACCCTAGTG CTACGGCCCACAGAAATGCCTTGAGAATTTTGGTCTCCTTCCTCTTGCTCT ATACTATATATTTCCTGTCTCTTCTTATATCATGGGTTGCTCAGAAGAATC AAAGTGAACTGGTTCACATTATTTGTATGATAACTTCACTCGTGTATCCTT CATTCCACTCATATATCCTGATTCTGGGAAATTATAAATTAAAGCAGACCT CTCTTTGGGTAATGAGGCAGCTGGGATGTAGGATGAAAAGACAGAATACAC CAACTACATAAGGCAGCCAAACAGTCTATTGGGTTTTAGATAACAAATCTA AATCTATGAGGAAGTAGTTCAATAACATTTTTCCCCTTGACATGGAGTAGC AGGGTTTTTTTTTATTAGATATTTTCTTTACTTACATTTCAAATGCTATCC CGAAAATTCCCTGTACCCTCTCCCTGTCCTGTTCCCCTACCCACCCACTCC CACTTCTTGGCCCTGGCATTCCCCTGGAGTATCAGTTTTTTATTAGTCAAA CTATCTCACTGACTAAGGGTCATAAAACAAGTTATTTTAACACTAATTTCA ATTAAATCAAAGGTAAAGTGTCAGCACATGCCTTTAATCACACAATTCCAT CAAATTCAGCACTCAGGAGAGGGTGATCTCTGTGAATTCCAGCACACTGGC GGCCGTTACTAGTGGATCCGAGCTCGGTACCAAGCTT
>mGR25 aa MMGIAIDILWAAIIIVQFIIGNIAN GFIALVNIIDWVKRRKISLMDKIIT ALAISRIYLLWSTFLITLTSSLDPD IKMAVKIIRISNNTWIIANHFSIWF ATCLSIFYFLKIANFSNYIFLYLRW RFKKVVSVTLLISLIFLLLNILLMN MHIDIWSDKSKRNLSFSVRSNNCTQ FPRLVLLINTMFTSIPFTVSLLAFL LLIFSLWRHLKTMQYYAKGSEDTTT AAHIKALHMVVAFLLFYTVFFLSLA IQYWTSGSQENNNLFYATIVITFPS VHSCILILRNSQLRQASLLVLWWLL CKSKDVRMLVP	>mGR25 nt AAAACTATTCGAATTGAACACAGTAACCAATTCTTCAGCGGACTTACACAA ATCAAGCTATTATCTTATGGATGATGGGTATTGCCATAGATATCTTATGGG CAGCTATTATCATTGTGCAATTCATAATTGGGAATATTGCAAATGGATTCA TAGCATTGGTGAACATCATAGACTGGGTGAAGAGAAGAAAAATCTCTTTAA TGGATAAGATCATTACTGCTTTGGCAATCTCTAGGATTTATCTGCTGTGGT CTACATTCTTAATTACACTAACATCTTCACTGGATCCAGATATTAAAATGG CTGTGAAAATCATTAGAATAAGCAATAACACCTGGATTATTGCAAATCATT TCAGCATTTGGTTTGCTACATGTCTCAGCATCTTTTATTTTCTCAAGATAG CCAATTTTTCTAACTATATTTTTCTCTACTTAAGGTGGAGATTTAAGAAGG TGGTTTCAGTGACATTGCTAATCTCTCTTATCTTCCTGCTTTTAAATATTT TACTGATGAACATGCATATTGATATCTGGAGTGATAAGTCCAAAAGAAACC TTTCTTTTAGTGTCAGATCAAATAATTGCACTCAGTTTCCCAGACTTGTCC TTTTAATCAACACAATGTTCACATCAATCCCCTTCACTGTGTCCCTGTTGG CTTTTCTGCTTCTCATCTTCTCCCTGTGGAGACACCTGAAAACCATGCAAT ACTATGCTAAAGGCTCCGAAGACACCACCACAGCTGCACATATAAAGGCCT TGCACATGGTAGTGGCCTTTCTCCTGTTCTACACAGTTTTCTTTTTGTCTC TTGCCATACAATATTGGACCTCTGGGTCTCAAGAGAATAACAACCTGTTTT ATGCCACAATTGTAATTACTTTCCCTTCAGTCCATTCATGTATCCTGATTC TGAGAAACAGCCAGCTGAGGCAGGCATCTCTGTTGGTGCTGTGGTGGCTGC TGTGCAAGTCCAAAGATGTACGGATGTTGGTTCCCTGAAATACTCTGTCAA TGCTCTTTAGTAGTGAAGAAGAAAATAGCTTAGTTAAGGAAATTCTTGTTC ATTACCGAAGTATACTTTCAAGTTTATGTATC
>mGR26 aa MLPTLSVFFMLTFVLLCFLGILANG FIVLMLSREWLLRGRLLPSDMILFS LGTSRFFQQCVGLVNSFYYFLHLVE YSGSLARQLISLHWDFLNSATFWFC	>mGR26 nt GAATTCTAGACAAGGAAAGACACACACTAAATGACTTTACTTGTGGGACCT AAAATAACCAAAATAAGTCAAAATCACAGTGATGTTACTAGGGATCTAGGA TAAGGGAATGAAGAGAAAGATGTTGGTCATAGAGTACAAAAATTCAGCTAA GAACTCAGTCCTGGAGGCTGAATGTATAGCTGTGTGACAGACAGCAGCTAG CCATACCAGAGTATACACTTGCCTCTTGCTGAAAGAGTAGATCTTATGTGT

TWLSVLFCIKIANFSH[redacted]LWLKWR FPALVPWFLLGSILVSVIVTLLFFW GNHTIYQAFLRRKFTGNTTFKEWNR RLEIDYFMPLKVVTMSIPCSLFLVS ILLLISSLRRHSLRMQHNTHSLQDP NVQAHSRALKSLISFLVLYAVSFVS MIIDATVFISSDNVWYWPWQIILYF CMSVHPFILITNNLRFRGTFRQLLL LARGFWVA	CCTTGTCACACATAAAAGTAATTGAAAAA[redacted]ACTCTCTGAGATGACAGAT ACGTTAAAATGGTTTTACTTTTCAACCTGC[redacted]CAGTAGGGGTCCCTTTAAT GTTTGTGCTAGTAGATGGGGGACTCTCAAGTATCTTTGTGGTAGACAAATC TAAGGTGGCCTTCATGAATACCAACCCAGACTTTTGTGACTTTGTGATCCC CCACTTTTGAAGTGGATAAGAGCTGTGACTTGAGTCTAATCAAAGGAGTCC AACGTGTTGTTTATTCTGTAACAGTGCTTTGTGTTTCTAGTTAATAACACA GGCAAAGAAGGCTAGGGTGACATTCCTAGGATTGTGTTATTTCTATCTTGC TCATGCCTCCCTCTGCTGGTCTAATGAAATAAGTCAGTGGCCATATTTAAA TATGACTACGTGGCAAATACTGATGATAGCCTGTGTGTTCCAACAAATATC CAGTAGGAGACCTAGGCATTCAGTCCTGCAGCCACAAGGAAATAGGTTCTT TCACTGGAAAAAGAGCAGTTTAGATGGTTATAAATTACTTAATCCATAGAA GCCATAGGGGCTTTATGTAGAGATTTGGGTAGAGAGGTAGACCTAGATATT GACTTAGGAGTGGCTATTCCTGAGTGGGGGTAGATATATGGCAGGGAAACT CAGATAAGAAAGACTTCTTTAGTGTCACGATTTTTCCTAGGTATCTCCTTG TGCCAGATATCTATGCGTCTATGTACCTACCTACCTACCTACCTACCTACC TACCTACCTACCTACTGACACCTAATAGGAAGAGGCAAGTGGTCACAACCT GCAATGATGGGATAAGAATGATGGAACTCAGTTACCAAGATTAAAATACCT TCCCCACTGATGTTATTGCAAGCATGGCAGCATGTAGGCAAAATCAGAGAA GGCAAATCATGAGCAGCTGCTGCCCCATGGTACCCGAGCCCGGGAAATATT TGCATCATATCTGAGCCAAAAGCACACCTTTTATCTACTGCCTGAGCATTT TTCACATTGAAGTTCTGGCTCACATGCAGAATCCAACCATTTATCTCCTGT CTCCAGAAGGGAGTGTCAGGGACTGTGGGTAGGGGCAGGGAGGAGGCCAGG AACCAAGGCAATCAGTGGTGACAGGAGGAGGGACTGAAATGCTACCAACAT TATCAGTTTTCTTCATGTTGACCTTTGTTCTGCTCTGTTTCCTGGGGATCC TGGCCAACGGCTTCATTGTGCTGATGCTGAGCAGGGAATGGCTACTGCGTG GTAGGCTGCTCCCCTCGGACATGATCCTCTTCAGTTTGGGCACCTCCCGAT TCTTCCAGCAGTGTGTGGGATTGGTCAACAGTTTCTATTACTTCCTCCATC TGGTTGAGTACTCCGGGAGCCTTGCCCGGCAGCTCATTAGTCTTCACTGGG ACTTCTTGAACTCAGCCACTTTCTGGTTTTGTACCTGGCTCAGCGTCCTGT TCTGTATCAAGATTGCTAACTTCTCCCATCCTGCCTTCCTGTGGTTGAAGT GGAGATTCCCAGCGTTGGTGCCCTGGTTCTTGTTGGGCTCTATCTTGGTGT CCGTCATTGTAACTCTGCTGTTCTTTTGGGGAAACCACACTATATATCAGG CATTCTTAAGGAGAAAGTTTACTGGGAACACAACCTTTAAGGAGTGGAACA GAAGGCTGGAAATAGACTATTTCATGCCTCTGAAAGTTGTCACCATGTCAA TTCCTTGTTCTCTTTTTCTGGTCTCAATTTTGCTGTTGATCAGTTCTCTCA GAAGGCATTCGCTAAGAATGCAGCACAATACCCACAGCTTGCAAGACCCCA ACGTCCAGGCTCACAGCAGAGCCCTGAAGTCACTCATCTCATTCCTGGTTC TTTATGCGGTGTCCTTTGTGTCCATGATCATTGATGCTACAGTCTTCATCT CCTCAGATAATGTGTGGTATTGGCCCTGGCAAATTATACTTTACTTTTGCA TGTCTGTACATCCATTTATCCTCATCACCAATAATCTCAGGTTCCGCGGCA CCTTCAGGCAGCTACTCCTGTTGGCCAGGGGATTCTGGGTGGCCTAGAAGG CTTGGTCTCTTTATCTAGAGCCTTTGAAGAGACTCAGGTGAGGGTAACTTC ACTTGGAAGTGAGCTCATCTACGTGGAAATGTCTTTGTAGGCAGGCATGGG GTCATACTGTGAGGTTCCTCATTGGGAAAGAGGAGAAGAAAATACAGAGTG TCCTTCCTTACCTTAGGATATTATGAAAGTGGAAATTCCGAATCCTGGACC AGTATTGATCTAAGTGCAAAGTACAATATGTCCTGTTCCTTTCATGTCTGT TTTCCTTTTGTTACTGATTCATTCTCTAGGGAATAGTCTTGATCAACTGAA TCATCTCATCTGGCTGGCCACTGGGGAGGTAAAAGAACTTTGTGTCACTGC TGCATTGGGATATACATGGGTGGGAAGCAAGTGTCCCTGAGGCAGAGTAGC ACTCAGTATGAGAACCTCAAAGAGCAGGTGGCTGTGCATGCAGGGGCTGGG GCAAGGAGTCCTGATCACTCTTCACTGTATGGGGATTATTTGTCTCTTGCC AAAATTTGGAGACTTTGGCTTTAGTTTTGTGAAGATGACTGGAAAAATTCT TAATGCTACCCTGTATCATTTCTCAATAATATTTTCCTTTTCCTGCCTTTA ATTTTCTCCTATCTGCAGCGCCCCTTGCTTGTTATCCGTAAATAAATAAAT AAATAAATAAATAAGCCCAATCCTCATTTTCCTGTCTTTGGGAACCCTTTT ACTTCCCCAGGTATACGCTACAAAGCCACTTCTGCATTGAATAAACATTAT CTTTCATTCAGAAAAAGACTTAAGAATCTCACCTTTACAAAAAAAAAAAAA AAAGAATCTCACTTATTTTATATTCAAATTCCATTTTTAAAAAGAAAAGCA CAGCATTAATTTTTCTAAATACTGTTTATAAAAATAACTTGCTCTAAGAAT TATACAAATGTTTTGAAAGGTAACTTTGGAAAAAAAGTGTGATTAGACATG GATGTTTGTAAGACAGAACAAAGAGCTCTTGGAAGTCCATGGCAGCTCATT GGTCTTGCCTTCAGTAGAGCCTGTCTGAATCCTGTAACCTCTTATGCCCTT TTGTAGCTTTTCTGCAGATC

>mGR27 aa	>mGR27 nt GAATTCGCCCTTGCGGGATCCGGGAACGGATTCATAGCACTGGTAAACTTC ATGGGCTGGATGAAGAATAGGAAGATTGCCTCCATTGATTTAATCCTCACA AGTCTGGCCATATCCAGAATTTGTCTATTGTGCGTAATACTATTAGATTGT TTTATATTGGTGCTATATCCAGATGTCTATGCCACTGGTAAAGAAATGAGA ATCATTGACTTCTTCTGGACACTAACCAATCACTTAAGTATCTGGTTTGCA ACCTGCCTCAGCATTTACTATTTCTTCAAGATAGGTAATTTCTTTCACCCA CTTTTCCTATGCCTCAAGTCTAGACGCCAAGGGC
>mGR28 aa GREWLRYGRLLPLDMILISLGASRF CLQLVGTVHNFYYSAQKVEYSGGLG RQFFHLHWHFLNSATFWFCSWLSVL FCVKIAN	>mGR28 nt GAATTCGCCCTTGCGGGATCCGGGAACGGGTTTATTGTGCTGGTGCTGGGC AGGGAGTGGCTGCGATATGGCAGGTTGCTGCCCTTGGATATGATCCTCATT AGCTTGGGTGCCTCCCGCTTCTGCCTGCAGTTGGTTGGGACGGTGCACAAC TTCTACTACTCTGCCCAGAAGGTCGAGTACTCTGGGGGTCTCGGCCGACAG TTCTTCCATCTACACTGGCACTTCCTGAACTCAGCCACCTTCTGGTTTTGC AGCTGGCTCAGTGTCCTGTTCTGTGTGAAGATTGCTAACATCACACACTCC ACCTTCCTGTGTCTCAAGTCTAGACGCCAAGGGCG
>mGR29 aa MDGIVQNMFTFIVIVEIIIGWIGNG FIALVNCIHWYKRRKISALNQILTA LAFSRIYLLLTVFTVIAVSTLYTHV LVTRRVVKLINFHLLFSNHFSMWLA ACLGLYYFLKIAHFPNSIFVYLKMR INQVVSGTLLMSLGLLFLNTLLINS YIDTKIDDYREHLLYDFTSNNTASF YRVILVINNCIFTSIPFTLSQSTFL LLIFSLWRHYKKMQQHAQRCRDVLA DAHIRVLQTMVTYVLLCAIFFLSLS MQILRSELLKNILYVRFCEIVAAVF PSGHSCVLICRDTNLRGTFLSVLSW LKQRFTSWIPNINCRSSCIF	>mGR29 nt AGCTTGATATTTCCTATTTGTTACTGCACAGAGTTTTTTTTAAAAATTGAG TTTGTTATGTGGATTCAATACTCAGATAGAGCTCTTTAATTTTTTTACAGT GACCTCATGAATCATAACTTGCCTTACAGACAATGGATGGAATCGTACAGA ACATGTTTACATTCATTGTAATTGTGGAAATAATAATAGGATGGATTGGAA ATGGATTCATAGCTCTGGTGAACTGCATACACTGGTACAAGAGAAGAAAGA TCTCTGCACTGAATCAAATACTCACAGCCTTGGCTTTCTCCAGAATCTACC TTCTTTTAACAGTATTCACTGTTATAGCAGTGTCTACGCTATACACACACG TGTTGGTAACTAGAAGAGTGGTAAAACTGATTAATTTCCATTTGCTTTTCA GCAATCATTTTAGCATGTGGCTTGCTGCATGCCTTGGCCTTTATTATTTTC TTAAAATAGCTCATTTTCCTAACTCTATTTTTGTTTACTTAAAGATGAGAA TTAACCAGGTGGTTTCAGGGACTTTGCTCATGTCTTTGGGCCTCTTGTTTC TAAACACTCTGCTGATAAACTCATACATTGATACCAAGATAGATGACTACA GAGAACATCTACTGTATGATTTCACTTCGAATAATACTGCTTCATTTTACA GGGTTATTTTAGTCATTAACAACTGTATTTTCACATCTATACCCTTTACAC TTTCCCAGTCCACTTTTCTCCTGCTCATCTTCTCCCTGTGGAGACATTACA AGAAGATGCAACAGCATGCACAAAGATGCAGAGATGTCCTTGCAGATGCCC ACATCAGAGTCTTGCAAACCATGGTCACCTATGTCCTACTCTGTGCCATTT TCTTTCTGTCTCTTTCCATGCAAATTTTGAGGAGTGAGTTGTTGAAGAACA TTCTTTACGTTAGGTTCTGCGAGATTGTTGCAGCAGTTTTTCCTTCAGGAC ACTCCTGTGTCTTAATCTGTAGAGACACAAACCTGAGAGGGACCTTTCTTT CTGTGCTATCGTGGCTGAAGCAGAGGTTTACATCATGGATTCCTAACATAA ATTGCAGATCATCTTGCATATTCTAAAAGAAACTGAG
>mGR30 aa MTYETDTTLMLVAVGEALVGILGNA FIALVNFMGWMKNRKIASIDLILSS VAMSRICLQCIILLDCIILVQYPDT YNRGKEMRTVDFFWTLTNHLSVWFA TCLSIFYLFKIANFFHPLFLWIKWR IDKLILRTLLACVIISLCFSLPVTE NLSDDFRRCVKTKERINSTLRCKVN KAGHASVKVNLNLVMLFPFSVSLVS FLLLILSLWRHTRQIQLSVTGYKDP STTAHVKAMKAVISFLALFVVYCLA FLIATSSYFMPESELAVIWGELIAL IYPSSHSFILILGSSKLKQASVRVL CRVKTMLKGKKY	>mGR30 nt AAAAATGTTCATTGTTTATCTAAAATTCAAATTTAACTGAGTGCCCTACAT TTTTATTTATTCAATCTAGTAGCTGTACTGAGGTTATTAGTGTGATTTCTG AAGCCCAAATTTGTAAAACTTAGCCTCAGATAAACAGCTTGAGACCATGGA AAGTAATTTGGTAAA*TTGCATCTTAGCAAA*TAGTAGCTCAGCCTAAATTA ACTGTGTGTAGAAAAGAATGACCTGCGGAGAAGATAAATGGACATACAATA TCCAGGCTAAGGATTGCCAAACACACTGTTTTTAAGACTAATTGAGATTTA GATAAACTATCTACAGTCTTCATGTATAATTCTCATCTTCATCACAAGACA GACTTCAACTTAAGGAGGTAAAGACAAGGACAGCGAACCCTAAACAGCCAA GTGTAGAAACCAAACTGCATCAAATCAGCCAGAAACTAATTGGATACTTCT CTACTTTAAAATGACATACGAAACAGATACTACCTTAATGCTTGTAGCTGT TGGTGAGGCCTTAGTAGGGATTTTAGGAAATGCATTCATTGCACTGGTAAA CTTCATGGGCTGGATGAAGAATAGGAAGATTGCCTCTATTGATTTAATCCT CTCAAGTGTGGCCATGTCCAGAATTTGTCTACAGTGTATAATCCTATTAGA TTGTATTATATTGGTGCAGTATCCAGACACCTACAACAGAGGTAAAGAAAT GAGGACCGTTGACTTCTTCTGGACACTTACCAACCATTTAAGTGTCTGGTT TGCCACCTGCCTCAGCATTTTCTATTTATTCAAGATAGCAAACTTCTTCCA CCCTCTTTTCCTCTGGATAAAGTGGAGAATTGACAAGCTAATTCTCAGAAC TCTACTGGCATGTGTGATTATCTCCCTGTGTTTTAGCCTCCCAGTCACTGA AAATCTGAGTGATGATTTCAGACGTTGTGTTAAGACAAAGGAGAGAATAAA CTCTACTTTGAGATGCAAAGTAAATAAAGCTGGACATGCCTCTGTCAAGGT AAATCTCAACTTGGTCATGCTGTTCCCCTTTTCTGTGTCTCTGGTCTCCTT TCTCCTCTTGATCCTCTCCCTGTGGAGACACACCAGGCAGATACAACTCAG TGTAACAGGGTACAAAGATCCCAGCACAACAGCTCATGTGAAAGCCATGAA

AGCAGTAATTTCCTTCCTGGCCCTGTTT[redacted]GTCTACTGCCTAGCCTTTCT
CATAGCCACCTCCAGCTACTTTATGCCAGAGAGTGAATTAGCTGTAATATG
GGGTGAGCTGATAGCTCTAATCTATCCTTCAAGCCATTCATTTATCCTCAT
CCTGGGGAGTAGTAAACTAAAACAAGCATCTGTGAGGGTGCTTTGTAGAGT
AAAGACCATGTTAAAGGGAAAAAAATATTAGCATCATGAGCATATCTGAAG
AAAAACTATCACTTTCTAAGAGAAAGGAAGACACGATCATTATCCGTCCTT
TTCACATGAATATTGATTTCATGCAGTGACATCCTCTTAACAAACTTAAAT
TGAACCTTGAGAAATCTCATATACAGCAACTTTGCATGTCTCTATCTCTGC
TTTTTCTCTCCTTTTCAATATGAGTTGACATAAAAAATAATTTTCAGAACA
AATTATAACAGAAGAAAGGGCATTTTCATAATCAGTTCTGAATCACTCCTC
CAAATGCAAAGCTGCCTGACAAATTCAAAACAATTGTAACAGCATCTCACT
GTCGTTTGCATTCTTTGGAAAAGCAGGTGGTTTGTTCTTGGAGCCTGGCTT
AGAGTTTTCTTCTTAGACCATTGAATTATGTTCATGATTGGAGAAGAGTCA
AGTACCAAGTAACAATTTTTATTGTGAAGATGGGTGTTCATCATGTGATTT
TGGCTGGCCTGGAACTTGTTATGTAGACTAGTCTGTCATCAAACACACAAA
GATCTGCCTGCCTCACCTGCCAGTTCTAGGATTCAAGGAATGCACCACCAC
AGCTTGTTCAAGTGACAATTCTTACAAATGTTTTAGAAATAAATAATATAC
TAGAAATTAACACTGAATGTAAGTGCTGTTTAGGTATAAATTATGATTAAA
TGTTATAGTTAGAAAATTATTTAAGATTATAGATCAGTGATGAAAATATTC
TAGAATAAGTTTTATGAAGAAACTTTTATAAAGAAACTGGAAAAAAATCTC
TTGATTGCATATTGAAACAAATTTCTCCAAAAAGAACACCTACAAATTTGC
TCTAGACATCTAGACTGTATCAAACAGTGAATATGAAAATATCATAACAGG
ATATAGCCTTTAGTATTGAAGACAGGTTCATCTATATTAAACCTGCATACA
TACCTAAAAGACTAAGTCAATATCCCACAAACATATTTGCACTATCATGTC
TATTGAAACACTATTCATAGTAGCTAAAATATGGCACAAAACTAGACATTC
ATCAATAGATGAATCAATAAAGCAAATGTACATACACAAGATGAAATTGTA
TTCAGGCATAAAGAAGAATGCAGTCATGTCATTAGCAAAAACATAAACAGA
ATTGGAGGTCATTGTGATAATTGAAATAAACCAGACCTGGAAAAAACAAAA
CCTGTGTAATTTTTCTGAAGTAGAGAATATACTCTTGGATGGATAGATGGG
TACTGTTATAGTATAAAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTAT
TTCATGAAAGCAAGAATGGGACTGCTTAGAGAAAGAAAAGGACAAACAGGT
GAAGGGGTGAAAGAAAAAGGCAATGACAAGGAGTAATGATATGAGCAAAGT
ACCATTATTAAACATGTGACAATATTATATAGAAACACATGATTTTGTGTG
CCTACCAAAACTGGATAATAATTTTTAAAATGTATCTATTAAAAGGAAAGA
AAAGAAAGTGCAAGCCCAGGAAAGGGAGAAAAGGAAACAATGAGAGAGAAA
TGGAAAATGGTGAGAAGTGAAGAGAACAAAAAGAAATGGAGTAAGTGTGGC
CAGGAATGAAGGATCTCAGCTATAGTTATCCCAGTACGGTAATACAAATCT
GTGACTCCAGCACTTGACAAGGCTGAGAGATGTGAGAGAGGGCCAGTTAAC
AACCAGTCTGGGCTTATTCCAAGAGATAAGAAGATTGGGGGAAAGTATGTA
GAAGGGTTTGGAGGGAAGAGAGAGAAGAGGGAAATGATGTAATGATAGTAC
AAATCAAAAGTTATTTTTTCTAAAAAAGCAATGGGACAGGAAACCAACCTA
ACAAGTAAAGGTGCTTGGTTCACAAGACCAGCAACCTGAGTGCATCCTTGC
TAGAATGAAATTGGCCTTACTCTGGAAAGCTTACTTCCTCAGTGTATTCAT
TGTTAAAATTCATGTGGAGATTTTAAAGAAAAAAGGAAAAAAAAAGTTAAA
TGGTAGATTTGTGTAGGGGAATATTCCCCTAATTAATTGATTAGATAATAA
AGATGACAAGCAAATTGCTGTGCAAAAAGGAAGACAAGGTCTAAGAGGGGA
AGAGGGGACACGGGAGGAAAAAAAACGGCCCTTTTTAAAGCAAGGTGGGGA
GTGAGGGAAGCGAGATGTAGACAGGGAACTGTTAGACCTGGTGGCAGCTTC
TGCCACCTGAAGATTTTCAACATAGTATAGTTCATGAGTTTAGGAAGATAT
GTTCCCTGCCCAGCGGTTGTATCATCTGTTGATTTTAAACTAAGATTGTCT
GGTGTTTTCCATTTGCGGAGACTCAAGTAGACCAAAGGGAAAGAATGAATT
C

>mGR31 aa
MYMILVRAVFITGMLGNMFIGLANC
SDWVKNQKITFINFIMVCLAASRIS
SVLMLFIDATIQELAPHFYYSYRLV
KCSDIFWVITDQLSTWLATCLSIFY
LFKVAHISHPLFLWLKWRLRGVLVV
FLVFSLFLLISYFLLLETLPIWGDI
YVTLKNNLTLFSGTIKTTAFQKIIV
FDIIYLVPFLVSLASLLLLFLSLVK
HSRSLDLISTTSEDSRTKIHKKAMK

>mGR31 nt
CTGCAGCTTTCTAGAAATCTCACCAGAATGTCTTTGTGCAGCTTTAATAGT
TCCTGGTTATACCTTGTCACATTATAAGCTAAGACATCTTTGGTGCCACAA
TATACTCTCACTAATCAGAGAGATTAGACAGAAAAATAAGTTTCTTAACAA
CTGTTTTAGATAGGGTCATGAAATGACATAAAACACCAATGCTAAGGCAAT
CCATTATGTTTTCTCATGAGGAGCCCATATGTACACTTGAGTGTGTCTTAT
TATTTCCCTGAGTGATTTTGTAATTTTATTAAACACTTAACTGTGATTCAT
ACTAGTTAGTTCTGAAATTCTTTTCTTCATCAAAGCCATTAATCCTGGGGT
TTTTTAAATGGAGAACCCCAAAACAAAGTGAAATGTTGTGTGTGGAGCAGG
CTGTCTTCCCACACACTACCATGAGATGCTCATTCTGTAATTGTTCCCCGG
AATAGGAAATGCCCTGAATTCAGGCACACAAGAGCTAGTCTGTGCACCATG

MLVSFLILFIIHIFF[redacted]LARWLLFL FPMSRPINFILTLNIFALTHSFILI LGNSNLRQRAMRILQHLKSQLQELI LSLHRFSSLY	TCTGGTTCTTGCATTAATACCCACTTTT[redacted]ACGAAGCTTCATTGATTCGC ATCTTCAGAAGCTGGTATCATTATTAGTTTCTTTCCTCAGGTGACTCTGGn CCAAAATATTAnGGCGCCCTTTAAAAAAGTAAAACTACAAAATTTCTTTAT AATTTTCTTTAAGTTTGTTATAATATAGCATGACCTACACACACACACACA CACACACACACACACACACACACACAAGTATGCCTCTCCTTTCCTTCTAAA AATCTCACTTAAAGCAATTGTTTAGCTGTCTTCGAAGTCTAGACTGCCACT GTCGTGCTTCTAGCCAAAACAAATGCAACACATAAAATGATAGAGCTCAAA ACTTAGGAATCTATTTAACTGTGAAGATCACGCAAGCAAACCTGAGAAACC TCTAGAAGGAAACCACAGCAAATCACTGGAGAGAAGGTGTTAATCTAGTAA GAATAGTTTTTATTTTGGGTATCCTTTTGTAGATTGGTTAGTTCATCCAAA ATCCAACTTGTTAGTTCTTCATAAATTGTAAGTGTCTCCAACATCAAAGCA CCACTTCTCTCTTTTCCCCTGTATGAAGATGCTTTAAGTACAGAGTTACTC TTTTTCTGTACTGACAGTAATTTAAAAAAATTGTTCACTCATTCTTTTTTG GTGTTGTTATTCTGTGTTCCTCAATGTTATCTTTTTTTTTTCAAAACTTTC TTTTATAAAAAGTCATACACATAGCAAATGCAGTGCATGTTTATGGAATCC ATAACTAACTTATTGAGACTTCTCCTAGTACTTTCTTTGAACAGTAACAAA GATATCTGCTTCTACAGAGTGCAGTGTTTCAGGTGAGGAGGAACATATTAT ACAAATCAGTGAAAAAAAAATCTGATTCAAATTTGTATTTTAATATATTTG ACTTTATCACTTCAGATATTACATCAATGGGAATTTTGAAGGCACACAAGT GATGATGTGGGCATAGAGACTGTCTGTACTAGAATTTAATATTTCTTTTAA ATATCTTTAAATAAAAATATGATGCTGTATTCATAAACAGATCTTTATAGA TTAAGTATGAGATTAAAGTTGGAAAAACAAAAGACAAAAACCTAGGACTAA GAATTTCCTTAAGTATGTGTGAATATCAACCTAATGGAGGAAGTTTCCAAT CAAAGCTGAAATTACAGTAAAAAGGAGGAAGATAAATATGGAAAAGGATGA TTTTCTGTGGAAGTTTGTTTGAGAACTGATCCACGAGACAAATTGCTAGAA GTGTGGATTCCCTTTTACTATTCAACTGCTTATAGGACTGGATCAAATGTA TATGATACTGGTAAGAGCAGTATTTATAACTGGAATGCTGGGAAATATGTT CATTGGACTGGCAAACTGCTCTGACTGGGTCAAGAACCAGAAAATCACCTT CATCAACTTCATCATGGTCTGTTTGGCAGCTTCCAGAATCAGCTCTGTGCT GATGTTATTTATTGATGCAACCATACAAGAACTAGCGCCTCATTTCTATTA TTCTTACCGTCTAGTAAAATGCTCTGATATATTCTGGGTTATAACTGATCA ACTATCAACATGGCTTGCCACCTGCCTGAGCATATTCTACTTATTCAAAGT AGCCCACATTTCCCATCCCCTTTTCCTCTGGTTGAAGTGGAGATTGAGAGG TGTGCTTGTTGTTTTTCTTGTATTTTCTTTGTTCTTATTGATTTCTTATTT TCTACTGCTTGAAACACTTCCTATTTGGGGAGATATTTATGTAACCCTTAA AAACAATCTGACCTTATTTTCAGGTACAATTAAGACCACTGCTTTTCAAAA GATAATTGTTTTTGATATAATATATTTAGTCCCATTTCTTGTGTCCCTAGC ATCATTGCTCCTTTTATTTTTGTCCTTGGTGAAACACTCCCGAAGCCTTGA CCTGATTTCTACCACTTCTGAAGATTCCAGAACCAAGATTCATAAGAAGGC CATGAAAATGCTGGTGTCTTTCCTCATTCTCTTTATAATTCACATTTTTTT CATGCAGTTAGCACGGTGGTTATTATTTTGTTTCCAATGAGCAGGCCAAT TAATTTCATCTTAACATTAAATATCTTTGCCTTAACTCACTCATTTATTCT CATCCTGGGAAATAGCAATCTTCGACAGAGAGCAATGAGGATCCTGCAACA TCTTAAAAGCCAGCTTCAAGAGCTGATCCTCTCCCTTCATAGATTCTCCAG TCTTTACTAGAGGAACAGCTTAACAGGGAGACTTGGAAGGTCACTGGCAAA TTATTCTTCTTTGATTTCTTTTAAGTACTGCTGAACATATATGAACTGTCC CCAGAGCATAGTGCTATCTTATGAGAAGGATATCATCTCACAGTCTGGTTA TAAAACACAAACCAATCTTTTTATAATTTCTTTACAGCATTGCTAATAAAA GACTTGTAGTCTCAAATATTTTAAAGAGAATAATTAATTTTATAGGCAAAA GGTATGAAATTACAATTCACAGGGAAGGTTCATGACTCCTTAGATATTAAA GTTAATTGTAAGCCACAATAGGCAGAAGATGAGCAAAATGTTGATAGGAGA TAAATAAAATCTAAAGTTACGGAGAAAAAAAAACATCAACTTGCCTTTTAGA TTACTTTAAAGCTCTCTCTCTCGGCTCTCTCTCTCTGTATCTACTTACTTTA TATATACAAATGTTTTGTCTGCATGTATTTCTTTGCACCATATAAATGTCT AAGTATCCAGAAnGTCAGCAGAGGGCATCAAATTCTCTGGAAAGAGAGTTA CAAATTGCTGTGGGTAACACTGGGTGCTGGGAACTAACCTGAGTCCTCTGC CACAGCAACTGCTCTTCCCTGCTGAGTCATGTTTTAAGTCTCCACAACTTA AACTCATTGTTGATGTGGTCATTGCATAATGATGAATTTACATTCTAAGGT TTGTATCATAGGTAGGAGGGCTGGTTTTAATCATATTCTAATGTTCTTATA CAAACCCAGGTTTTGTAAGAGACTGTATTCTATCATGAGACTCTTTCCCCA CACCGCCAATGTAACATTTTTATTAATTTTGAGGGGAATTTTATACAGTGT ACCCTGATCACCCTTGCTTCCCACTCCTTGCAGGTCTACCCTCCCACCATT GCTCAATCCCCCCTAAAAGAGAGAGAAACAAACCATGTCCAATTTGTGTTG

	GACACATACTCAGTGGAACATGGCCAAA[redacted]TAGTGAGCAGTTCCTTAAA GAAAACTAAGCTGCCTCCCCACCACTACC[redacted]ATAGGGCATTAACTGTGAA GAGCTACACTTTAGCTATTTTATCACCAATTTAAAAGACTGTCTTCAATAG CTTCCTCTATGGACTGTTTCTGGTTTTAGTGGGACAGGGAGAAGGGGTCAA GAGGTTGTCACAGAAACTTTTGATGTCTCTTATTCTCAGTTAAAGTCCACT GCAAAAGAAGTCTGCTGGCTCTAATAAAGCTTGCAACAGCATGGGCCAGTG ACATCATCATGATTTCTGGCAACAATATGGACCACAAATATCATGGCTCAG GTGGCATTACGGACCACAGACATCAACATGGTCTCTGGCAGCAAGAACCAG AATCTTTTGAGGAGGCTTCATTCAGAAAATGAATTTTTCTTCATCCCAGAT ATACTGATGTTGCTCAATCAGAGTATTAGTATGGTTGGGCACCATATTTGG GGACAGGACCTTCAATATTTCCAGGCTGCTGTGTAACACATTATCTTTAGT GTCAGGTGCCCTTAGTGTCAGGACATGACCATCATGTATGCGCCTGTGGGC AGAAATACATCTTTGTACTTTCTTACACCTAGCAGGGTGAGTAGCAGGAGC AGCGGCATTAATACTTCCATACCTCTGGGCAGCCTATCAGGTATCATCTAG GCAAGGTAAGCCCAGTAGTGGCCCAAGGCTCCTGGTGTCTACTTGGCAACA ACATGCTCCTTTGTCTGCACTGCCATATCTATGGCTGGTTCTCCATCCCTA GTTCTGCTTCTCTCAGGTTTTATACGACTCTATTCCACATTCTATTTTTCC AGTTCCATGAAACCAGTGTTTAAAAGTATCATCCCATAAGACCGGCCTTTT AAAGGTTATTCTGGAGATATTGCAGAGTCTGCAG

SEQUENCE LISTING

SEQ ID NO:1

Human T2R01 amino acid sequence

MLESHLIIYFLLAVIQFLLGIFTNGIIVVVNGIDLIKHRKMAPLDLLLSCLAVSRIFLQL
FIFYVNVIVIFFIEFIMCSANCAILLFINELELWLATWLGVFYCAKVASVRHPLFIWLKM
RISKLVPWMILGSLLYVSMICVFHSKYAGFMVPYFLRKFFSQNATIQKEDTLAIQIFSFV
AEFSVPLLIFLFAVLLLIFSLGRHTRQMRNTVAGSRVPGRGAPISALLSILSFLILYFSH
CMIKVFLSSLKFHIRRFIFLFFILVIGIYPSGHSLILILGNPKLKQNAKKFLLHSKCCQ

SEQ ID NO:2

Human T2R01 nucleotide sequence

ATGCTAGAGTCTCACCTCATTATCTATTTTCTTCTTGCAGTGATACAATTTCTTCTTGGG
ATTTTCACAAATGGCATCATTGTGGTGGTGAATGGCATTGACTTGATCAAGCACAGAAAA
ATGGCTCCGCTGGATCTCCTTCTTTCTTGTCTGGCAGTTTCTAGAATTTTTCTGCAGTTG
TTCATCTTCTACGTTAATGTGATTGTTATCTTCTTCATAGAATTCATCATGTGTTCTGCG
AATTGTGCAATTCTCTTATTTATAAATGAATTGGAACTTTGGCTTGCCACATGGCTCGGC
GTTTTCTATTGTGCCAAGGTTGCCAGCGTCCGTCACCCACTCTTCATCTGGTTGAAGATG
AGGATATCCAAGCTGGTCCCATGGATGATCCTGGGGTCTCTGCTATATGTATCTATGATT
TGTGTTTTCCATAGCAAATATGCAGGGTTTATGGTCCCATACTTCCTAAGGAAATTTTTC
TCCCAAAATGCCACAATTCAAAAAGAAGATACACTGGCTATACAGATTTTCTCTTTTGTT
GCTGAGTTCTCAGTGCCATTGCTTATCTTCCTTTTTGCTGTTTTGCTCTTGATTTTCTCT
CTGGGGAGGCACACCCGGCAAATGAGAAACACAGTGGCCGGCAGCAGGGTTCCTGGCAGG
GGTGCACCCATCAGCGCGTTGCTGTCTATCCTGTCCTTCCTGATCCTCTACTTCTCCCAC
TGCATGATAAAAGTTTTTCTCTCTTCTCTAAAGTTTCACATCAGAAGGTTCATCTTTCTG
TTCTTCATCCTTGTGATTGGTATATACCCTTCTGGACACTCTCTCATCTTAATTTTAGGA
AATCCTAAATTGAAACAAAATGCAAAAAAGTTCCTCCTCCA**CAGTAAGTGCTGTCAGTGA**

SEQ ID NO:3

Human T2R02 amino acid sequence

MALSFSAILHIIMMSAEFFTGITVNGFLIIVNCNELIKHRKLMPIQILLMCIGMSRFGLQ
MVLMVQSFFSVFFPLLYVKIIYGAAMMFLWMFFSSISLWFATCLSVFYCLKISGFTQSCF
LWLKFRIPKLIPWLFWEAFWPL*ALHLCVEVDYAKNVEEDALRNTTLKKSKTKIKKISEV
LLVNLALIFPLAIFVMCTSMLLISLYKHTHRMQHGSHGFRNANTEAHINALKTVITFFCF
FISYFAAFMTNMTFSLPYRSHQFFMLKDIMAAYPSGHSVIIILSNSKFQQSFRRILCLKK
KL

SEQ ID NO:4

Human T2R02 nucleotide sequence

ATGGCCTTGTCTTTTTCAGCTATTCTTCATATTATCATGATGTCAGCAGAATTCTTCACA
GGGATCACAGTAAATGGATTTCTTATCATTGTTAACTGTAATGAATTGATCAAACATAGA
AAGCTAATGCCAATTCAAATCCTCTTAATGTGCATAGGGATGTCTAGATTTGGTCTGCAG
ATGGTGTTAATGGTACAAAGTTTTTTCTCTGTGTTCTTTCCACTCCTTTACGTCAAAATA
ATTTATGGTGCAGCAATGATGTTCCTTTGGATGTTTTTTAGCTCTATCAGCCTATGGTTT
GCCACTTGCCTTTCTGTATTTTACTGCCTCAAGATTTCAGGCTTCACTCAGTCCTGTTTT
CTTTGGTTGAAATTCAGGATCCCAAAGTTAATACCTTGGCTGCTTCTGGGAAGCGTTCTG
GCCTCTGTGAGCATTGCATCTGTGTGTCGAGGTAGATTACGCTAAAAATGTGGAAGAGGA
TGCCCTCAGAAACACCACACTAAAAAAGAGTAAAACAAAGATAAAGAAAATTAGTGAAGT
GCTTCTTGTCAACTTGGCATTAATATTTCCTCTAGCCATATTTGTGATGTGCACTTCTAT
GTTACTCATCTCTCTTTACAAGCACACTCATCGGATGCAACATGGATCTCATGGCTTTAG
AAATGCCAACACAGAAGCCCATATAAATGCATTAAAAACAGTGATAACATTCTTTTGCTT
CTTTATTTCTTATTTTGCTGCCTTCATGACAAATATGACATTTAGTTTACCTTACAGAAG
TCACCAGTTCTTTATGCTGAAGGACATAATGGCAGCATATCCCTCTGGCCACTCGGTTAT
AATAATCTTGAGTAATTCTAAGTTCCAACAATCATTTAGAAGAATTCTCTG**CCTCAAAAA**
GAAACTATGA

SEQ ID NO:5

Human T2R03 amino acid sequence

MMGLTEGVFLILS●QFTLGILVNCFIELVNGSSWFKTKRMSLS●IITTLALLRIILLC
IILTDSFLIEFSPNTHDSGIIMQIIDVSWTFTNHLSIWLATCLGVLYCLKIASFSHPTFL
WLKWRVSRVMVWMLLGALLLSCGSTASLINEFKLYSVFRGIEATRNVTEHFRKKRSEYYL
IHVLGTLWYLPPLIVSLASYSLLIFSLGRHTRQMLQNGTSSRDPTTEAHKRAIRIILSFF
FLFLLYFLAFLIASFGNFLPKTKMAKMIGEVMTMFYPAGHSFILILGNSKLKQTFVVMLR
CESGHLKPGSKGPIFS

SEQ ID NO:6

Human T2R03 nucleotide sequence

ATGATGGGACTCACCGAGGGGGTGTTCCTGATTCTGTCTGGCACTCAGTTCACACTGGGA
ATTCTGGTCAATTGTTTCATTGAGTTGGTCAATGGTAGCAGCTGGTTCAAGACCAAGAGA
ATGTCTTTGTCTGACTTCATCATCACCACCCTGGCACTCTTGAGGATCATTCTGCTGTGT
ATTATCTTGACTGATAGTTTTTTAATAGAATTCTCTCCCAACACACATGATTCAGGGATA
ATAATGCAAATTATTGATGTTTCCTGGACATTTACAAACCATCTGAGCATTTGGCTTGCC
ACCTGTCTTGGTGTCCTCTACTGCCTGAAAATCGCCAGTTTCTCTCACCCCACATTCCTC
TGGCTCAAGTGGAGAGTTTCTAGGGTGATGGTATGGATGCTGTTGGGTGCACTGCTCTTA
TCCTGTGGTAGTACCGCATCTCTGATCAATGAGTTTAAGCTCTATTCTGTCTTTAGGGGA
ATTGAGGCCACCAGGAATGTGACTGAACACTTCAGAAAGAAGAGGAGTGAGTATTATCTG
ATCCATGTTCTTGGGACTCTGTGGTACCTGCCTCCCTTAATTGTGTCCCTGGCCTCCTAC
TCTTTGCTCATCTTCTCCCTGGGGAGGCACACACGGCAGATGCTGCAAAATGGGACAAGC
TCCAGAGATCCAACCACTGAGGCCCACAAGAGGGCCATCAGAATCATCCTTTCCTTCTTC
TTTCTCTTCTTACTTTACTTTCTTGCTTTCTTAATTGCATCATTTGGTAATTTCCTACCA
AAAACCAAGATGGCTAAGATGATTGGCGAAGTAATGACAATGTTTTATCCTGCTGGCCAC
TCATTTATTCTCATTCTGGGGAACAGTAAGCTGAAGCAGACATTTGTAGTGATGCTCCGG
TGTGAGTCTGGTCATCTGAAGCCTGGATCCAA**GGGACCCATTTTCTCTTAG**

SEQ ID NO:7

Human T2R04 amino acid sequence

MLRLFYFSAIIASVILNFVGIIMNLFITVVNCKTWVKSHRISSSDRILFSLGITRFLMLG
LFLVNTIYFVSSNTERSVYLSAFFVLCFMFLDSSSVWFVTLLNILYCVKITNFQHSVFLL

LKRNISPKIPRLI●CVLISAFTTCLYITLSQASPFPELVTTRN●SFNISEGILSLVVS
LVLSSSLQFIINVTSASLLIHSLRRHIQKMQKNATGFWNPQTEAHVGAMKLMVYFLILYI
PYSVATLVQYLPFYAGMDMGTKSICLIFATLYSPGHSVLIIITHPKLKTTAKKILCFKK

SEQ ID NO:8

Human T2R04 nucleotide sequence

ATGCTTCGGTTATTCTATTTCTCTGCTATTATTGCCTCAGTTATTTTAAATTTTGTAGGA
ATCATTATGAATCTGTTTATTACAGTGGTCAATTGCAAAACTTGGGTCAAAAGCCATAGA
ATCTCCTCTTCTGATAGGATTCTGTTCAGCCTGGGCATCACCAGGTTTCTTATGCTGGGA
CTATTTCTGGTGAACACCATCTACTTCGTCTCTTCAAATACGGAAAGGTCAGTCTACCTG
TCTGCTTTTTTTGTGTTGTGTTTCATGTTTTTGGACTCGAGCAGTGTCTGGTTTGTGACC
TTGCTCAATATCTTGTACTGTGTGAAGATTACTAACTTCCAACACTCAGTGTTTCTCCTG
CTGAAGCGGAATATCTCCCCAAAGATCCCCAGGCTGCTGCTGGCCTGTGTGCTGATTTCT
GCTTTCACCACTTGCCTGTACATCACGCTTAGCCAGGCATCACCTTTTCCTGAACTTGTG
ACTACGAGAAATAACACATCATTTAATATCAGTGAGGGCATCTTGTCTTTAGTGGTTTCT
TTGGTCTTGAGCTCATCTCTCCAGTTCATCATTAATGTGACTTCTGCTTCCTTGCTAATA
CACTCCTTGAGGAGACATATACAGAAGATGCAGAAAAATGCCACTGGTTTCTGGAATCCC
CAGACGGAAGCTCATGTAGGTGCTATGAAGCTGATGGTCTATTTCCTCATCCTCTACATT
CCATATTCAGTTGCTACCCTGGTCCAGTATCTCCCCTTTTATGCAGGGATGGATATGGGG
ACCAAATCCATTTGTCTGATTTTTGCCACCCTTTACTCTCCAGGACATTCTGTTCTCATT
ATTATCACACATCCTAAACTGAAAACAACAGCAAA**GAAGATTCTTTGTTTCAAAAAATAG**

SEQ ID NO:9

Human T2R05 amino acid sequence

MLSAGLGLLMLVAVVEFLIGLIGNGSLVVWSFREWIRKFNWSSYNLIILGLAGCRFLLQW
LIILDLSLFPLFQSSRWLRYLSIFWVLVSQASLWFATFLSVFYCKKITTFDRPAYLWLKQ
RAYNLSLWCLLGYFIINLLLTVQIGLTFYHPPQGNSSIRYPFESWQYLYAFQLNSGSYLP
LVVFLVSSGMLIVSLYTHHKKMKVHSAGRRDVRAKAHITALKSLGCFLLLHLVYIMASPF
SITSKTYPPDLTSVFIWETLMAAYPSLHSLILIMGIPRVKQTCQKILWKTVCARRCWGP

SEQ ID NO:10

Human T2R05 nucleotide sequence

ATGCTGAGCGCTGGCCTAGGACTGCTGATGCTGGTGGCAGTGGTTGAATTTCTCATCGGT
TTAATTGGAAATGGAAGCCTGGTGGTCTGGAGTTTTAGAGAATGGATCAGAAAATTCAAC
TGGTCCTCATATAACCTCATTATCCTGGGCCTGGCTGGCTGCCGATTTCTCCTGCAGTGG
CTGATCATTTTGGACTTAAGCTTGTTTCCACTTTTCCAGAGCAGCCGTTGGCTTCGCTAT
CTTAGTATCTTCTGGGTCCTGGTAAGCCAGGCCAGCTTATGGTTTGCCACCTTCCTCAGT
GTCTTCTATTGCAAGAAGATCACGACCTTCGATCGCCCGGCCTACTTGTGGCTGAAGCAG
AGGGCCTATAACCTGAGTCTCTGGTGCCTTCTGGGCTACTTTATAATCAATTTGTTACTT
ACAGTCCAAATTGGCTTAACATTCTATCATCCTCCCCAAGGAAACAGCAGCATTCGGTAT
CCCTTTGAAAGCTGGCAGTACCTGTATGCATTTCAGCTCAATTCAGGAAGTTATTTGCCT
TTAGTGGTGTTTCTTGTTTCCTCTGGGATGCTGATTGTCTCTTTGTATACACACCACAAG
AAGATGAAGGTCCATTCAGCTGGTAGGAGGGATGTCCGGGCCAAGGCTCACATCACTGCG
CTGAAGTCCTTGGGCTGCTTCCTCTTACTTCACCTGGTTTATATCATGGCCAGCCCCTTC
TCCATCACCTCCAAGACTTATCCTCCTGATCTCACCAGTGTCTTCATCTGGGAGACACTC
ATGGCAGCCTATCCTTCTCTTCATTCTCTCATATTGATCATGGGGATTCCTAGGGTGAAG
CAGACTTGTCAGAAGATCCTGTGGAAGACAGTGTGTGCTCG**GAGATGCTGGGGCCCATGA**

SEQ ID NO:11

Human T2R06 amino acid sequence

MLAAALGLLMPIAGAEFLIGLVGNGVPVVCSFRGWVKKM*GVPINSHDSGK*PLSPTQAD
HVGHKSVSTFPEQWLALLS*CLRVLVSQANM*FATFFSGFCCMEIMTFVXXXXXXXXXXX
XXXXXXXXXLLVSFKITFYFSALVGWTL*KPLTGNSNILHPILNLLFL*IAVQ*RRLIAI
CDVSVPLVFL*RHHRKMEDHTAVRRRLKPRXXXXXXXXXXXXXXXLYMVSALARHFSMTF
*SPSDLTILAISATLMAVYTSFPSIVMVMRNQTCQRIL*EMICTWKS

SEQ ID NO:12

Human T2R06 nucleotide sequence

ATGTTGGCGGCT●CCTAGGATTGCTGATGCCCATTGCAGGGG●AATTTCTCATTGGC
CTGGTTGGAAATGGAGTCCCTGTGGTCTGCAGTTTTAGAGGATGGGTCAAAAAAATGTAA
GGAGTCCCTATAAATTCTCATGATTCTGGTAAGTAGCCACTTTCTCCTACTCAGGCCGAT
CATGTTGGACATAAGTCTGTTTCCACTTTCCCAGAGCAGTGGTTGGCTTTACTATCTTAA
TGTCTTCGAGTCCTGGTAAGCCAGGCCAACATGTAGTTTGCCACTTTCTTCAGTGGCTTC
TGCTGCATGGAGATCATGACCTTTGTCCCGCTGACTTCTTGTAGCTGAAAAGACTGGGTT
TTTGTTTTTTGCTAGTGTCTTTCAAGATCACTTTTTATTTCTCAGCTCTTGTTGGCTGGA
CCCTTTAAAAACCCTTAACAGGAAACAGCAACATCCTGCATCCCATTTTAAATCTGTTAT
TTTTATAGATTGCTGTCCAGTGAAGGAGACTGATTGCTATTTGTGATGTTTCTGTTCCAC
TTGTCTTTTTGTAAAGACATCACAGGAAGATGGAGGACCACACAGCTGTCAGGAGGAGGC
TCAAACCAAGGTGCTCATCGCTCTGAACTTCCCCCTTTACATGGTTTCTGCCTTGGCCAG
ACACTTTTCCATGACCTTCTAATCTCCCTCTGATCTCACCATTCTTGCCATCTCTGCAAC
ACTCATGGCTGTTTATACTTCATTTCCGTCTATTGTAATGGTTATGAGGAATCAGACTTG
TCAGAGAATTCTGTAGGAGATGATATGTACATGGAAATCCTAG

SEQ ID NO:13

Human T2R07 amino acid sequence

MADKVQTTLLFLAVGEFSVGILGNAFIGLVNCMDWVKKRKIASIDLILTSLAISRICLLC
VILLDCFILVLYPDVYATGKEMRIIDFFWTLTNHLSIWFATCLSIYYFFKIGNFFHPLFL
WMKWRIDRVISWILLGCVVLSVFISLPATENLNADFRFCVKAKRKTNLTWSCRVNKTQHA
STKLFLNLATLLPFCVCLMSFFLLILSLRRHIRRMQLSATGCRDPSTEAHVRALKAVISF
LLLFIAYYLSFLIATSSYFMPETELAVIFGESIALIYPSSHSFILILGNNKLRHASLKVI
WKVMSILKGRKFQQHKQI

SEQ ID NO:14

Human T2R07 nucleotide sequence

ATGGCAGATAAAGTGCAGACTACTTTATTGTTCTTAGCAGTTGGAGAGTTTTCAGTGGGG
ATCTTAGGGAATGCATTCATTGGATTGGTAAACTGCATGGACTGGGTCAAGAAGAGGAAA
ATTGCCTCCATTGATTTAATCCTCACAAGTCTGGCCATATCCAGAATTTGTCTATTGTGC
GTAATACTATTAGATTGTTTTATATTGGTGCTATATCCAGATGTCTATGCCACTGGTAAA

GAAATGAGAATCA●GACTTCTTCTGGACACTAACCAATCATTT●GTATCTGGTTTGCA
ACCTGCCTCAGCATTTACTATTTCTTCAAGATAGGTAATTTCTTTCACCCACTTTTCCTC
TGGATGAAGTGGAGAATTGACAGGGTGATTTCCTGGATTCTACTGGGGTGCGTGGTTCTC
TCTGTGTTTATTAGCCTTCCAGCCACTGAGAATTTGAACGCTGATTTCAGGTTTTGTGTG
AAGGCAAAGAGGAAAACAAACTTAACTTGGAGTTGCAGAGTAAATAAAACTCAACATGCT
TCTACCAAGTTATTTCTCAACCTGGCAACGCTGCTCCCCTTTTGTGTGTGCCTAATGTCC
TTTTTCCTCTTGATCCTCTCCCTGCGGAGACATATCAGGCGAATGCAGCTCAGTGCCACA
GGGTGCAGAGACCCCAGCACAGAAGCCCATGTGAGAGCCCTGAAAGCTGTCATTTCCTTC
CTTCTCCTCTTTATTGCCTACTATTTGTCCTTTCTCATTGCCACCTCCAGCTACTTTATG
CCAGAGACGGAATTAGCTGTGATTTTTGGTGAGTCCATAGCTCTAATCTACCCCTCAAGT
CATTCATTTATCCTAATACTGGGGAACAATAAATTAAGACATGCATCTCTAAAGGTGATT
TGGAAAGTAATGTCTATTCTAAAAGGAAGAAAATT**CCAACAACATAAACAAATCTGA**

SEQ ID NO:15

Human T2R08 amino acid sequence

MFSPADNIFIILITGEFILGILGNGYIALVNWIDWIKKKKISTVDYILTNLVIARICLIS
VMVVNGIVIVLNPDVYTKNKQQIVIFTFWTFANYLNMWITTCLNVFYFLKIASSSHPLFL
WLKWKIDMVVHWILLGCFAISLLVSLIAAIVLSCDYRFHAIAKHKRNITEMFHVSKIPYF
EPLTLFNLFAIVPFIVSLISFFLLVRSLWRHTKQIKLYATGSRDPSTEVHVRAIKTMTSF
IFFFFLYYISSILMTFSYLMTKYKLAVEFGEIAAILYPLGHSLILIVLNNKLRQTFVRML
TCRKIACMI

SEQ ID NO:16

Human T2R08 nucleotide sequence

ATGTTCAGTCCTGCAGATAACATCTTTATAATCCTAATAACTGGAGAATTCATACTAGGA
ATATTGGGGAATGGATACATTGCACTAGTCAACTGGATTGACTGGATTAAGAAGAAAAAG
ATTTCCACAGTTGACTACATCCTTACCAATTTAGTTATCGCCAGAATTTGTTTGATCAGT
GTAATGGTTGTAAATGGCATTGTAATAGTACTGAACCCAGATGTTTATACAAAAAATAAA
CAACAGATAGTCATTTTTACCTTCTGGACATTTGCCAACTACTTAAATATGTGGATTACC
ACCTGCCTTAATGTCTTCTATTTTCTGAAGATAGCCAGTTCCTCTCATCCACTTTTTCTC

TGGCTGAAGTGG[redacted]AATTGATATGGTGGTGCACTGGATCCTGCT[redacted]GATGCTTTGCCATT
TCCTTGTTGGTCAGCCTTATAGCAGCAATAGTACTGAGTTGTGATTATAGGTTTCATGCA
ATTGCCAAACATAAAAGAAACATTACTGAAATGTTCCATGTGAGTAAAATACCATACTTT
GAACCCTTGACTCTCTTTAACCTGTTTGCAATTGTCCCATTTATTGTGTCACTGATATCA
TTTTTCCTTTTAGTAAGATCTTTATGGAGACATACCAAGCAAATAAAACTCTATGCTACC
GGCAGTAGAGACCCCAGCACAGAAGTTCATGTGAGAGCCATTAAAACTATGACTTCATTT
ATCTTCTTTTTTTTCCTATACTATATTTCTTCTATTTTGATGACCTTTAGCTATCTTATG
ACAAAATACAAGTTAGCTGTGGAGTTTGGAGAGATTGCAGCAATTCTCTACCCCTTGGGT
CACTCACTTATTTTAATTGTTTTAAATAATAAACTGAGGCAGACATTTGTCAGAATGCTG
ACATGTA**GAAAAATTGCCTGCATGATATGA**

SEQ ID NO:17

Human T2R09 amino acid sequence

MPSAIEAIYIILIAGELTIGIWGNGFIVLVNCIDWLKRRDISLIDIILISLAISRICLLC
VISLDGFFMLLFPGTYGNSVLVSIVNVVWTFANNSSLWFTSCLSIFYLLKIANISHPFFF
WLKLKINKVMLAILLGSFLISLIISVPKNDDMWYHLFKVSHEENITWKFKVSKIPGTFKQ
LTLNLGVMVPFILCLISFFLLLFSLVRHTKQIRLHATGFRDPSTEAHMRAIKAVIIFLLL
LIVYYPVFLVMTSSALIPQGKLVLMIGDIVTVIFPSSHSFILIMGNSKLREAFLKMLRFV
KCFLRRRKPFVP

SEQ ID NO:18

Human T2R09 nucleotide sequence

ATGCCAAGTGCAATAGAGGCAATATATATTATTTTAATTGCTGGTGAATTGACCATAGGG
ATTTGGGGAAATGGATTCATTGTACTAGTTAACTGCATTGACTGGCTCAAAAGAAGAGAT
ATTTCCTTGATTGACATCATCCTGATCAGCTTGGCCATCTCCAGAATCTGTCTGCTGTGT
GTAATATCATTAGATGGCTTCTTTATGCTGCTCTTTCCAGGTACATATGGCAATAGCGTG
CTAGTAAGCATTGTGAATGTTGTCTGGACATTTGCCAATAATTCAAGTCTCTGGTTTACT
TCTTGCCTCAGTATCTTCTATTTACTCAAGATAGCCAATATATCGCACCCATTTTTCTTC
TGGCTGAAGCTAAAGATCAACAAGGTCATGCTTGCGATTCTTCTGGGGTCCTTTCTTATC
TCTTTAATTATTAGTGTTCCAAAGAATGATGATATGTGGTATCACCTTTTCAAAGTCAGT

CATGAAGAAAAC[redacted]TACTTGGAAATTCAAAGTGAGTAAAATTC[redacted]GTACTTTCAAACAG
TTAACCCTGAACCTGGGGGTGATGGTTCCCTTTATCCTTTGCCTGATCTCATTTTTCTTG
TTACTTTTCTCCCTAGTTAGACACACCAAGCAGATTCGACTGCATGCTACAGGGTTCAGA
GACCCCAGTACAGAGGCCCACATGAGGGCCATAAAGGCAGTGATCATCTTTCTGCTCCTC
CTCATCGTGTACTACCCAGTCTTTCTTGTTATGACCTCTAGCGCTCTGATTCCTCAGGGA
AAATTAGTGTTGATGATTGGTGACATAGTAACTGTCATTTTCCCATCAAGCCATTCATTC
ATTCTAATTATGGGAAATAGCAAGTTGAGGGAAGCTTTTCTGAAGATGTTAAGATTTGTG
AAGTGTTTCCTTAGAAGAA**GAAAGCCTTTTGTTCCATAG**

SEQ ID NO:19

Human T2R10 amino acid sequence

MLRVVEGIFIFVVVSESVFGVLGNGFIGLVNCIDCAKNKLSTIGFILTGLAISRIFLIWI
IITDGFIQIFSPNIYASGNLIEYISYFWVIGNQSSMWFATSLSIFYFLKIANFSNYIFLW
LKSRTNMVLPFMIVFLLISSLLNFAYIAKILNDYKTKNDTVWDLNMYKSEYFIKQILLNL
GVIFFFTLSLITCIFLIISLWRHNRQMQSNVTGLRDSNTEAHVKAMKVLISFIILFILYF
IGMAIEISCFTVRENKLLLMFGMTTTAIYPWGHSFILILGNSKLKQASLRVLQQLKCCEK
RKNLRVT

SEQ ID NO:20

Human T2R10 nucleotide sequence

ATGCTACGTGTAGTGGAAGGCATCTTCATTTTTGTTGTAGTTAGTGAGTCAGTGTTTGGG
GTTTTGGGGAATGGATTTATTGGACTTGTAAACTGCATTGACTGTGCCAAGAATAAGTTA
TCTACGATTGGCTTTATTCTCACCGGCTTAGCTATTTCAAGAATTTTTCTGATATGGATA
ATAATTACAGATGGATTTATACAGATATTCTCTCCAAATATATATGCCTCCGGTAACCTA
ATTGAATATATTAGTTACTTTTGGGTAATTGGTAATCAATCAAGTATGTGGTTTGCCACC
AGCCTCAGCATCTTCTATTTCCTGAAGATAGCAAATTTTTCCAACTACATATTTCTCTGG
TTGAAGAGCAGAACAAATATGGTTCTTCCCTTCATGATAGTATTCTTACTTATTTCATCG
TTACTTAATTTTGCATACATTGCGAAGATTCTTAATGATTATAAAACGAAGAATGACACA
GTCTGGGATCTCAACATGTATAAAAGTGAATACTTTATTAAACAGATTTTGCTAAATCTG
GGAGTCATTTTCTTCTTTACACTATCCCTAATTACATGTATTTTTTTAATCATTTCCCTT

TGGAGACACAACA[redacted]CAGATGCAATCGAATGTGACAGGATTGAG[redacted]ACTCCAACACAGAA
GCTCATGTGAAGGCAATGAAAGTTTTGATATCTTTCATCATCCTCTTTATCTTGTATTTT
ATAGGCATGGCCATAGAAATATCATGTTTTACTGTGCGAGAAAACAAACTGCTGCTTATG
TTTGGAATGACAACCACAGCCATCTATCCCTGGGGTCACTCATTTATCTTAATTCTAGGA
AACAGCAAGCTAAAGCAAGCCTCTTTGAGGGTACTGCAGCAATTGAAGTGCTGTGAGAAA
AG**GAAAAATCTCAGAGTCACATAG**

SEQ ID NO:21

Human T2R11 amino acid sequence

MANMLKNMLTMISAIDFIMGIQRSRVMVLVHCIDWIRRWKLSLIDFILTCWAISRIFXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXNHLCT*FATCLAVFYFLKIVNFSYLFYFWLK
WRINKVAFILPLVSAFSVYQLSFDVHF*CLLVSCPKKYERHMTGLLNVSNNKNVNNIIIF
FIGSLSSFSISSIFFLLLLLSS*RHMKHIRFNFRDCRTPVYGPISEPRKRFSFFVLLLYK
NLPFS

SEQ ID NO:22

Human T2R12 amino acid sequence

MSSIWETLFIRILVV*FIMGTVGN*FIVLVNIID*IRN*KVSLIDFILNCLAISRICFL*
ITILATSFNIGYEKMPDSKNLAVSFDILWTGSSYFCLSCTTCLSVFYFLKVANFSNPIFL
WMKWKIHKVLLFIVLEATISFCTTSILKEIIINSLI*ERVTIKGNLTFNYMDTMHDFTSL
FLLQMMFILPFVETLASILLLILSLWSHTRQMKLHGIYSRDPSTEAHVKPIKAIISFLLL
FIVHYFISIILTLACPLLDFVAARTFSSVLVFFHPSGHSFLLILRDSKLKQASLCVLKKM
KYAKKDIISHFYKHA

SEQ ID NO:23

Human T2R12 nucleotide sequence

ATGTCAAGCATTTGGGAGACACTGTTTATAAGAATTCTTGTAGTGTAATTCATAATGGGG
ACTGTGGGAAATTGATTCATTGTATTGGTTAATATCATTGACTGAATCAGGAACTGAAAG

GTCTCCCTGATTC[redacted]TTTATTCTCAACTGCTTGGCCATCTCCAG[redacted]TATGTTTCCTGTAG
ATAACAATTTTAGCTACCTCTTTCAATATAGGCTATGAGAAAATGCCTGATTCTAAGAAT
CTTGCAGTAAGTTTTGACATTCTCTGGACAGGATCCAGCTATTTCTGCCTGTCCTGTACC
ACTTGCCTCAGTGTCTTCTATTTCCTCAAGGTAGCCAACTTCTCCAATCCCATTTTCCTC
TGGATGAAATGGAAAATTCACAAGGTGCTTCTCTTTATTGTACTAGAGGCAACGATCTCT
TTCTGCACAACTTCCATTCTGAAGGAAATAATAATTAATAGTTTAATCTAAGAACGGGTA
ACAATAAAAGGCAACTTGACATTTAATTATATGGATACCATGCATGATTTCACTTCTCTG
TTTCTCCTTC**AG**ATGATGTTCATCCTTCCTTTTGTGGAAACACTGGCTTCCATTCTTCTC
TTAATCCTCTCCTTATGGAGCCACACCAGGCAGATGAAGCTACATGGTATTTATTCCAGG
GATCCCAGCACAGAAGCCCATGTAAAACCTATAAAAGCTATAATTTCATTTCTACTCCTC
TTTATTGTGCATTATTTCATCAGTATCATACTAACATTGGCCTGTCCTCTTCTAGACTTC
GTTGCGGCAAGGACTTTTAGTAGTGTGCTGGTATTTTTCCATCCATCTGGCCATTCATTT
CTTCTAATTTTACGGGACAGCAAACTGAAGCAAGCTTCTCTCTGTGTCCTGAAGAAGATG
AAGTATGCCAAAAAGGACATAATCT**CTCATTTTTATAAACATGCCTGA**

SEQ ID NO:24

Human T2R13 amino acid sequence

MESALPSIFTLVIIAEFIIGNLSNGFIVLINCIDWVSKRELSSVDKLLIILAISRIGLIW
EILVSWFLALHYLAIFVSGTGLRIMIFSWIVSNHFNLWLATIFSIFYLLKIASFSSPAFL
YLKWRVNKVILMILLGTLVFLFLNLIQINMHIKDWLDRYERNTTWNFSMSDFETFSVSVK
FTMTMFSLTPFTVAFISFLLLIFSLQKHLQKMQLNYKGHRDPRTKVHTNALKIVISFLLF
YASFFLCVLISWISELYQNTVIYMLCETIGVFSPSSHSFLLILGNAKLRQAFLLVAAKVW
AKR

SEQ ID NO:25

Human T2R13 nucleotide sequence

ATGGAAAGTGCCCTGCCGAGTATCTTCACTCTTGTAATAATTGCAGAATTCATAATTGGG
AATTTGAGCAATGGATTTATAGTACTGATCAACTGCATTGACTGGGTCAGTAAAAGAGAG
CTGTCCTCAGTCGATAAACTCCTCATTATCTTGGCAATCTCCAGAATTGGGCTGATCTGG
GAAATATTAGTAAGTTGGTTTTTAGCTCTGCATTATCTAGCCATATTTGTGTCTGGAACA

GGATTAAGAATTA[redacted]ATTTTTAGCTGGATAGTTTCTAATCACTT[redacted]ATCTCTGGCTTGCT
ACAATCTTCAGCATCTTTTATTTGCTCAAAATAGCGAGTTTCTCTAGCCCTGCTTTTCTC
TATTTGAAGTGGAGAGTAAACAAAGTGATTCTGATGATACTGCTAGGAACCTTGGTCTTC
TTATTTTTAAATCTGATACAAATAAACATGCATATAAAAGACTGGCTGGACCGATATGAA
AGAAACACAACTTGGAATTTCAGTATGAGTGACTTTGAAACATTTTCAGTGTCGGTCAAA
TTCACTATGACTATGTTCAGTCTAACACCATTTACTGTGGCCTTCATCTCTTTTCTCCTG
TTAATTTTCTCCCTGCAGAAACATCTCCAGAAAATGCAACTCAATTACAAAGGACACAGA
GACCCCAGGACCAAGGTCCATACAAATGCCTTGAAAATTGTGATCTCATTCCTTTTATTC
TATGCTAGTTTCTTTCTATGTGTTCTCATATCATGGATTTCTGAGCTGTATCAGAACACA
GTGATCTACATGCTTTGTGAGACGATTGGAGTCTTCTCTCCTTCAAGCCACTCCTTTCTT
CTGATTCTAGGAAACGCTAAGTTAAGACAGGCCTTTCTTTTGGTGGCAGCTAA**GGTATGG**
GCTAAACGATGA

SEQ ID NO:26

Human T2R14 amino acid sequence

MGGVIKSIFTFVLIVEFIIGNLGNSFIALVNCIDWVKGRKISSVDRILTALAISRISLVW
LIFGSWCVSVFFPALFATEKMFRMLTNIWTVINHFSVWLATGLGTFYFLKIANFSNSIFL
YLKWRVKKVVLVLLLVTSVFLFLNIALINIHINASINGYRRNKTCSSDSSNFTRFSSLIV
LTSTVFIFIPFTLSLAMFLLLIFSMWKHRKKMQHTVKISGDASTKAHRGVKSVITFFLLY
AIFSLSFFISVWTSERLEENLIILSQVMGMAYPSCHSCVLILGNKKLRQASLSVLLWLRY
MFKDGEPSGHKEFRESS

SEQ ID NO:27

Human T2R14 nucleotide sequence

ATGGGTGGTGTCATAAAGAGCATATTTACATTCGTTTTAATTGTGGAATTTATAATTGGA
AATTTAGGAAATAGTTTCATAGCACTGGTGAACTGTATTGACTGGGTCAAGGGAAGAAAG
ATCTCTTCGGTTGATCGGATCCTCACTGCTTTGGCAATCTCTCGAATTAGCCTGGTTTGG
TTAATATTCGGAAGCTGGTGTGTGTCTGTGTTTTTCCCAGCTTTATTTGCCACTGAAAAA
ATGTTCAGAATGCTTACTAATATCTGGACAGTGATCAATCATTTTAGTGTCTGGTTAGCT
ACAGGCCTCGGTACTTTTTATTTTCTCAAGATAGCCAATTTTTCTAACTCTATTTTTCTC

TACCTAAAGTGGA[redacted]GGTTAAAAAGGTGGTTTTGGTGCTGCTTCT[redacted]TGACTTCGGTCTTC
TTGTTTTTAAATATTGCACTGATAAACATCCATATAAATGCCAGTATCAATGGATACAGA
AGAAACAAGACTTGCAGTTCTGATTCAAGTAACTTTACACGATTTTCCAGTCTTATTGTA
TTAACCAGCACTGTGTTCATTTTCATACCCTTTACTTTGTCCCTGGCAATGTTTCTTCTC
CTCATCTTCTCCATGTGGAAACATCGCAAGAAGATGCAGCACACTGTCAAAATATCCGGA
GACGCCAGCACCAAAGCCCACAGAGGAGTTAAAAGTGTGATCACTTTCTTCCTACTCTAT
GCCATTTTCTCTCTGTCTTTTTTCATATCAGTTTGGACCTCTGAAAGGTTGGAGGAAAAT
CTAATTATTCTTTCCCAGGTGATGGGAATGGCTTATCCTTCATGTCACTCATGTGTTCTG
ATTCTTGGAAACAAGAAGCTGAGACAGGCCTCTCTGTCAGTGCTACTGTGGCTGAGGTAC
ATGTTCAAAGATGGGGAGCCCTCAGGTCACAAA**GAATTTAGAGAATCATCTTGA**

SEQ ID NO:28

Human T2R15 amino acid sequence

MITFLPIIFSILVVVTFVLGNFANGFIVLVNSIEWVKRQKISFADQILTALAVSRVGLLW
VILLHWYATVLNPGSYSLGVRITTINAWAVTNHFSIWVATSLSIFYFLKIANFSNFIFLH
LKRRIKSVIPVILLGSLLFLVCHLVVVNMDESMWTKEYEGNVSWEIKLSDPTHLSDMTVT
TLANLIPFTLSLLSFLLLICSLCKHLKKMQFHGKGSPDSNTKVHIKALQTVTSFLLLFAV
YFLSLITSIWNFRRRL*NEPVLMLSQTTAIIYPSFHSFILIWGSKKLKQTFLLILCQIKC

SEQ ID NO:29

Human T2R15 nucleotide sequence

ATGATAACTTTTCTACCCATCATTTTTTCCATTCTAGTAGTGGTTACATTTGTTCTTGGG
AATTTTGCTAATGGCTTCATAGTGTTGGTAAATTCCATTGAGTGGGTCAAGAGACAAAAG
ATCTCCTTTGCTGACCAAATTCTCACTGCTCTGGCAGTCTCCAGAGTTGGTTTGCTCTGG
GTAATATTATTACATTGGTATGCAACTGTTTTGAATCCAGGTTCATATAGTTTAGGAGTA
AGAATTACTACTATTAATGCCTGGGCTGTAACCAACCATTTCAGCATCTGGGTTGCTACT
AGCCTCAGCATATTTTATTTCCTCAAGATTGCCAATTTCTCCAACTTTATTTTTCTTCAC
TTAAAAAGGAGAATTAAGAGTGTCATTCCAGTGATACTATTGGGGTCTTTGTTATTTTTG
GTTTGTCATCTTGTTGTGGTAAACATGGATGAGAGTATGTGGACAAAAGAATATGAAGGA
AACGTGAGTTGGGAGATCAAATTGAGTGATCCGACGCACCTTTCAGATATGACTGTAACC

ACGCTTGCAAACT[redacted]ATACCCTTTACTCTGTCCCTGTTATCTTT[redacted]TGCTCTTAATCTGT
TCTTTGTGTAAACATCTCAAGAAGATGCAGTTCCATGGCAAAGGATCTCCAGATTCCAAC
ACCAAGGTCCACATAAAAGCTTTGCAAACGGTGACCTCCTTCCTCTTGTTATTTGCTGTT
TACTTTCTGTCCCTAATCACATCGATTTGGAATTTTAGGAGGAGGCTGTAGAACGAACCT
GTCCTCATGCTCAGCCAAACTACTGCAATTATATACCCTTCATTTCATTCATTCATCCTA
ATTTGGGGAAGCAAGAAGCTGAAACAGACCTTTCTTTTGATTTT**GTGTCAGATTAAGTGC**
TGA

SEQ ID NO:30

Human T2R16 amino acid sequence

MIPIQLTVFFMIIYVLESLTIIVQSSLIVAVLGREWLQVRRLMPVDMILISLGISRFCLQ
WASMLNNFCSYFNLNYVLCNLTITWEFFNILTFWLNSLLTVFYCIKVSSFTHHIFLWLRW
RILRLFPWILLGSLMITCVTIIPSAIGNYIQIQLLTMEHLPRNSTVTDKLENFHQYQFQA
HTVALVIPFILFLASTIFLMASLTKQIQHHSTGHCNPSMKARFTALRSLAVLFIVFTSYF
LTILITIIGTLFDKRCWLWVWEAFVYAFILMHSTSLMLSSPTLKRILKGKC

SEQ ID NO:31

Human T2R16 nucleotide sequence

ATGATACCCATCCAACTCACTGTCTTCTTCATGATCATCTATGTGCTTGAGTCCTTGACA
ATTATTGTGCAGAGCAGCCTAATTGTTGCAGTGCTGGGCAGAGAATGGCTGCAAGTCAGA
AGGCTGATGCCTGTGGACATGATTCTCATCAGCCTGGGCATCTCTCGCTTCTGTCTACAG
TGGGCATCAATGCTGAACAATTTTTGCTCCTATTTTAATTTGAATTATGTACTTTGCAAC
TTAACAATCACCTGGGAATTTTTTAATATCCTTACATTCTGGTTAAACAGCTTGCTTACC
GTGTTCTACTGCATCAAGGTCTCTTCTTTCACCCATCACATCTTTCTCTGGCTGAGGTGG
AGAATTTTGAGGTTGTTTCCCTGGATATTACTGGGTTCTCTGATGATTACTTGTGTAACA
ATCATCCCTTCAGCTATTGGGAATTACATTCAAATTCAGTTACTCACCATGGAGCATCTA
CCAAGAAACAGCACTGTAACTGACAAACTTGAAAATTTTCATCAGTATCAGTTCCAGGCT
CATACAGTTGCATTGGTTATTCCTTTCATCCTGTTCCTGGCCTCCACCATCTTTCTCATG
GCATCACTGACCAAGCAGATACAACATCATAGCACTGGTCACTGCAATCCAAGCATGAAA
GCGCGCTTCACTGCCCTGAGGTCCCTTGCCGTCTTATTTATTGTGTTTACCTCTTACTTT

CTAACCATACTC[redacted]CACCATTATAGGTACTCTATTTGATAAGAC[redacted]GTTGGTTATGGGTC
TGGGAAGCTTTTGTCTATGCTTTCATCTTAATGCATTCCACTTCACTGATGCTGAGCAGC
CCTACGTTGAAAAG**GATTCTAAAGGGAAAGTGCTAG**

SEQ ID NO:32

Human T2R17 amino acid sequence

MCSAXLLIILSILVVFAFVLGNVANGFIALINVNDWVKTQKISSTDQIVTALAFSRIGLL
XTLIILLHWYATVFNSALYSLEVRIVPSNVSAIINHFSIWLATSLSIFYLFKIANFSNFI
FLHLKKRIKSVLLVILLGSLVFLICNLAVVTMDDSVWTKEFEGNVTWKIELRNAIHLSNM
TITNHASKLHTVHSDSNIFSAVSLFSXTMLANFTLFILTLISFLLLVCSPCKHLKMMQLH
GKGSQDLSTKVHIKPLQTVISFRMLFAIYFLCIITSTWNPRTQQSNLVFLLYQTLAIMYP
SFHSFILIMRSRKLKQTSLSVLCQVTCWVK

SEQ ID NO:33

Human T2R18 amino acid sequence

MFVGINIFFLVVATRGLVLGMLGNGLIGLVNCIEWAKSWKVSSADFILTSLAIVRIIRLY
LILFDSFIMVLSPHLYTIRKLVKLFTILWALINQLSI*FATCLSIFYLLKIANFSHSLFL
WLKWRMNGMIVMLLILSLFLLIFDSLVLEIFIDISLNIIDKSNLTLYLDESKTLYDKLSI
LKTLLSLTYVIPFLLTLTSLLLLFISLVRHTKNLQLNSLGSRDSSTEAHKRAMKMVIAFL
LLFIINFISTLIGDWIFLEVENYQVMMFIMMILLAFPSGHSFIIILGNNKLRQSSLRLLW
HLKFSLKKAKPLTS

SEQ ID NO:34

Human T2R18 nucleotide sequence

ATGTTCGTTGGAATTAATATTTTCTTTCTGGTGGTGGCAACAAGAGGACTTGTCTTAGGA
ATGCTGGGAAACGGGCTCATTGGACTGGTAAACTGCATTGAGTGGGCCAAGAGTTGGAAG
GTCTCATCAGCTGATTTCATCCTCACCAGCTTGGCTATAGTCAGAATCATTCGACTGTAT
TTAATACTATTTGATTCATTTATAATGGTATTGTCCCCTCATCTATATACCATCCGTAAA

CTAGTAAAACTGT[redacted]ACTATTCTTTGGGCATTAATTAATCAGTT[redacted]GTATCTAGTTTGCC
ACCTGCCTAAGCATTTTCTACTTGCTTAAGATAGCCAATTTCTCCCACTCCCTTTTCCTC
TGGCTGAAGTGGAGAATGAACGGAATGATTGTTATGCTTCTTATATTGTCTTTGTTCTTA
CTGATTTTTGACAGTTTAGTGCTAGAAATATTTATTGATATCTCACTCAATATAATAGAT
AAAAGTAATCTGACTTTATATTTAGATGAAAGTAAAACTCTCTATGATAAACTCTCTATT
TTAAAAACTCTTCTCAGCTTGACATACGTTATTCCCTTTCTTCTGACTCTGACCTCTTTG
CTCCTTTTATTTATATCCTTAGTGAGACACACCAAGAATTTGCAGCTCAACTCTCTGGGC
TCAAGGGACTCCAGCACAGAGGCCCATAAAAGGGCCATGAAAATGGTGATAGCCTTCCTC
CTCCTTTTTATTATTAACTTTATTTCCACTTTAATAGGAGATTGGATCTTCCTTGAGGTA
GAGAATTATCAGGTCATGATGTTTATTATGATGATTTTACTTGCCTTTCCCTCAGGCCAC
TCATTTATTATAATTTTGGGAAACAACAAGCTAAGACAGAGCTCCTTGAGACTACTGTGG
CATCTTAAATTCTCTCTGAAAAAAGCAAAACCTTTAACTTCATAG

SEQ ID NO:35

Human T2R19 amino acid sequence

VTTLANLIPFTLSLICFLLLICSLCKHLKKMRLHSKGSQDPSTKVHIKALQTVTSFLMLF
AIYFLCIITSTWNLRTQQSKLVLLLCQTVAIMYPSFHSFILIMGSRKLKQTFLSVLWQMT
C

SEQ ID NO:36

Human T2R19 nucleotide sequence

CTGTAACTACTCTAGCAAACCTCATACCCTTTACTCTGAGCCTAATATGTTTTCTGCTGT
TAATCTGTTCTCTTTGTAAACATCTCAAGAAGATGCGGCTCCATAGCAAAGGATCTCAAG
ATCCCAGCACCAAGGTCCATATAAAAGCTTTGCAAACTGTGACCTCCTTCCTCATGTTAT
TTGCCATTTACTTTCTGTGTATAATCACATCAACTTGGAATCTTAGGACACAGCAGAGCA
AACTTGTACTCCTGCTTTGCCAAACTGTTGCAATCATGTATCCTTCATTCCACTCATTCA
TCCTGATTATGGGAAGTAGGAAGCTAAAACAGACCTTTCTTTCAGTTTTGTGGCAGATGA
CATGCTGAGTGAAAGAAGAGAAACCCTCAACTCCATAGATTCACAAGGGGAGCATCGTGG
GTCTTCTAGCAGAAAACAAACTGATGGTGTCTGGAACATTTTATAT

SEQ ID NO:37

Human T2R20 amino acid sequence

HLXRKAKSVVLVIVLGSLFFLVCQLVMKNTYINVWTEECEGNVTWKIKLRNAMHLSNLTV
AMLANLIPFTLTVISFLLLIYSLCKHLKKMQLHGKGSQDPSTKIHIKALQTVTSFLVLLA
IYFLCLIIS

SEQ ID NO:38

Human T2R20 nucleotide sequence

TTCATCACTTAN**AAAGGAAGGCTAAGAGTGTAG**TTCTGGTGATAGTGTTGGGGTCTTTGT
TCTTTTTGGTTTGTCAACTTGTGATGAAAAACACGTATATAAATGTGTGGACAGAAGAAT
GTGAAGGAAACGTAACTTGGAAGATCAAACTGAGGAATGCAATGCACCTTTCCAACTTGA
CTGTAGCCATGCTAGCAAACTTGATACCATTCACTCTGACCGTGATATCTTTTCTGCTGT
TAATCTACTCTCTGTGTAAACATCTGAAGAAGATGCAGCTCCATGGCAAAGGATCTCAAG
ATCCCAGCACCAAGATCCACATAAAAGCTCTGCAAACTGTGACCTCCTTCCTCGTATTAC
TTGCCATTTACTTTCT**GTGTCTAATCATATCCTTTTG**

SEQ ID NO:39

Human T2R21 amino acid sequence

MPPGIGNTFLIVMMGEFII*MLGNGFIVLVNCIDW*GVK*SY*TTASSPAWLSPQSVNFG
*YYLIHL*QHYGHIYMPSIN**NLFIFFGH*PIT*LPGLLP*CFLLL*NTYFSHPCFIWL
RWRISRTLLELPLGSLLLLFFNLALTGGLSDLWINIYTIYERNSTWSLDVSKILYCSLWI
LVSLIYLISFLLSLISLLLLILSLMRHIRNLQLNTMGPRDLRMKAHKRAMKMKMKMMVSF
LLFFLVHFSSLLPTGWIFLIQQK*QANFFVLLTSIIFPSSHSFVLILENCKLRQTAVGPL
WHLKCHLKRVKL

SEQ ID NO:40

Human T2R22 amino acid sequence

MATESDTNLLILAIAEFIISMLGNVFIGLVNCSEXIKNXKVFSADFILTCLAISHNGQLL
VILFDSFLVGLASHLYTTYRLXKNCIMLWT

SEQ ID NO:41

Human T2R22 nucleotide sequence

TATAGGGACN**GTGATGCTTCGTACACTCTC**CAAGAAGAAACACTCCGTGAGGTATGTGAG
ACTGCATNCCTTAGTAGATCTNTTGGGATATATATTCATAATATAGAAAAANAGGCAAAG
ACTTNCTTAAGTATATGAGACTCTATCCAACAGCAGAAGGTTCTGATCAAGACTGGAAGT
GCAATANAAGCAATGAAGATAAGTATCAGATATGAATGCTCTTCTGCAATGGTCTGATTG
TNACATTATTAATGATACANAGTATTAAAAACTTGGATTTTNTTGTCTCTGGAGATGGCC
ACCGAATCGGACACAAATCTTCTGATTCTGGCAATAGCAGAATTCATCATCAGCATGCTG
GGGAATGTGTTCATTGGACTGGTAAACTGCTCTGAANGGATCAAGAACCANAAGGTCTTC
TCAGCTGACTTCATCCTCACCTGCTTGGCTATCTCTCACAATGGACAACTGTTGGTGATA
CTGTTTGATTCATTTCTAGTGGGACTTGCTTCACATCTATATACCACATATAGACTANGA
AAAAACTGTATTATGCTTTGGACATGACTAATCACTTGACACACTGCTTCGCACGTGCTA
GCATATTCTATTCTTAGATAGCCACTTCNCACTCCTT**GTCTCTGCTGAAGTGGGAT**

SEQ ID NO:42

Human T2R23 amino acid sequence

VAFVLGNVANGFIALVNVIDXVNTRKISSAEQILTALVVSRIGXTLXHSIP*DATRC*SA
LYRXEVRIVASN

SEQ ID NO:43

Human T2R23 nucleotide sequence

AGGGTTGAGTCGTGCTTATCTTCACTTAACCTAGTATANAANTACAGCATATAGCAAGGA
GAGAATGTATATGAAGAGGAGTGAATTTGAGTCTGTTTGAGAATAATGACCTTTTCTATT
TCTATAAAGACAGTTTTGAATTCATCTATTAGCATATGCTGGTGCTTGCCTGTTGACACT

AGTCACTGAATT[redacted]AAGGCAGAAAATGTTATTGCACATTTAGT[redacted]TCAAGTGTTCATCGA
AGTTAACATCTGGATGTTAAAGGACTCAGAACAAGTGTTACTAAGCCTGCATTTTTTTAT
CTGTTCAAACATGATGTGTTNTCTGCTCATCATTTCATCAATTCTGGTAGAGTTGCATTT
GTTCTTGGAAATGTNGCCAATGGCTTCATAGCTCTAGTAAATGTCATTGACTGNGTTAAC
ACACGAAAGATCTCCTCAGCTGAGCAAATTCTCACTGCTCTGGTGGTCTCCAGAATTGGT
NNTACTCTGNGTCATAGTATTCCTTGAGATGCAACTAGATGTTAATCTGCTCTATATAGG
NTAGAAGTAAGAATTGTTGCTTCTAATGCCTGAGCTCGTACGAACCATT

SEQ ID NO:44

Human T2R24 amino acid sequence

MATELDKIFLILAIAEFIISMLGNVFIGLVNCSEGIKNQKVFSADFILTCLAISTIGQLL
VILFDSFLVGLASHLYTTYRLGKTVIMLWHMTNHLTTWLATCLSIFYFFKIAHFPHSLFL
WLRWRMNGMIVMLLILSLFLLIFDSLVLEIFIDISLNIIDKSNLTLYLDESKTLYDKLSI
LKTLLSLTSFIPFSLFLTSLLFLFLSLVRHTRNLKLSSLGSRDSSTEAHRRAMKMVMSFL
FLFIVHFFSLQVANGIFFMLWNNKYIKFVMLALNAFPSCHSFILILGNSKLRQTAVRLLW
HLRNYTKTPNALPL

SEQ ID NO:45

Human T2R24 nucleotide sequence

ATGGCCACCGAATTGGACAAAATCTTTCTGATTCTGGCAATAGCAGAATTCATCATCAGC
ATGCTGGGGAATGTGTTCATTGGACTGGTAAACTGCTCTGAAGGGATCAAGAACCAAAAG
GTCTTCTCAGCTGACTTCATCCTCACCTGCTTGGCTATCTCCACAATTGGACAACTGTTG
GTGATACTGTTTGATTCATTTCTAGTGGGACTTGCTTCACATTTATATACCACATATAGA
CTAGGAAAAACTGTTATTATGCTTTGGCACATGACTAATCACTTGACAACCTGGCTTGCC
ACCTGCCTAAGCATTTTCTATTTCTTTAAGATAGCCCACTTCCCCCACTCCCTTTTCCTC
TGGCTGAGGTGGAGGATGAACGGAATGATTGTTATGCTTCTTATATTGTCTTTGTTCTTA
CTGATTTTTGACAGTTTAGTGCTAGAAATATTTATTGATATCTCACTCAATATAATAGAT
AAAAGTAATCTGACTTTATATTTAGATGAAAGTAAAACTCTCTATGATAAACTCTCTATT
TTAAAAACTCTTCTCAGCTTAACCAGTTTTATCCCCTTTTCTCTGTTCCTGACCTCCTTG
CTTTTTTTATTTCTGTCCTTGGTGAGACATACTAGAAATTTGAAGCTCAGTTCCTTGGGC

TCTAGAGACTCC[redacted]CACAGAGGCCCATAGGAGGGCCATGAAAAT[redacted]TGATGTCTTTCCTT
TTCCTCTTCATAGTTCATTTTTTTTCCTTACAAGTGGCCAATGGGATATTTTTTATGTTG
TGGAACAACAAGTACATAAAGTTTGTCATGTTAGCCTTAAATGCCTTTCCCTCGTGCCAC
TCATTTATTCTCATTCTGGGAAACAGCAAGCTGCGACAGACAGCTGTGAGGCTACTGTGG
CATCTTAGGAACTATACAAAAACACCAAATGCTTTACCTTTGTAG

SEQ ID NO:46

Human T2R25 amino acid sequence

LSPFRMLFAIYFLCIITSTWNPRTQQSNLVFLLYQTLAIMYPSFHSFILIMRSRKLKQTS
LSVLCQVTCWVK

SEQ ID NO:47

Human T2R26 amino acid sequence

MPPGIGNTFLIVMMGEFII*MLGNGFIVLVNCIDVRSQMILLDNCILTSLAISTISQLWI
ILLDSFVTALWPHLYAFNKLIKFIHIFWALTNHLVTWLACCLSVFYFFKIAYFSHPCFIW
LRWRISRTLLELPLGSLLLLFFNLALTGGLSDLWINIYTMYERNSTWSLDVSKILYCSLW
ILVSLIYLISFLLSLISLLLLILSLMRHIRNLQLNTMGPRDLRMKAHKRAMKMKMKMMVS
FLLFFLVHFSSLLPTGWIFLIQQK

SEQ ID NO:48

Human T2R27 amino acid sequence

LANLIDWAENQICLMDFILSSLAICRTLLLGCCVAIRCTYNDYPNIDAVNHNLIKIITIF
DILRLVSK*LGIWFASYLSIFYLLKVALFHHAIFLWLKWRISRAVFTFLMIFLFFYISII
SMIKIKLFLDQC*YKI*EKLLLEGRCE*SPPSC*PDAH*PGVVYSLYHFSYLMFLVCYLP
KGKHCTAVVIGDWLQRPRTEAYVRAMNIMIAFFFHLLYSLGTSLSSVSYFLCKRKIVALG
AYLSYPLSHSFILIMENNKVRKAL

SEQ ID NO:49

Human T2R28 amino acid sequence

NICVLLIILSILVVSAFVLGNVANGFIALINVNDW

SEQ ID NO:50

Human T2R29 amino acid sequence

MQAALTAFFVLLFSLLSLLGIAANGFIVLVLGKEWL

SEQ ID NO:51

Human T2R30 amino acid sequence

MITFLPIIFSILVVVTFVLGNFSNGFIALVNSIEWVKTRKISSADQILTALVVSRVGLLW
VILLHWYANVFNSALYSSEVGAVASNISAIINHFSIWLATSLSIFYLLKIANFSNLIFLH
LKKRIRSVVLVILLGPLVFLICNLAVITMDDSVWTKEYEGNVTWKIKLRNAIHLSNMTVS
TLANLIPFILTLICFLLLICSLCKHLKKMQLHGKGSQDPSTKVHIKALQTVTSFLLLCAI
YFLSMIISVCNFGRLEKQPVFMFCQAIIFSYPSTHPFILILGNKKLKQIFLSVLRHVRYW
VKDRSLRLHRFTRGALCVF

SEQ ID NO:52

Human T2R30 nucleotide sequence

ATGATAACTTTTCTACCCATCATTTTTTCCATTCTGGTAGTGGTTACATTTGTTCTTGGA
AATTTTTCCAATGGCTTCATAGCTCTAGTAAATTCCATTGAGTGGGTCAAGACACGAAAG
ATCTCCTCAGCTGACCAAATCCTCACTGCTCTGGTGGTCTCCAGAGTTGGTTTACTCTGG
GTCATATTATTACATTGGTATGCAAATGTGTTTAATTCAGCTTTATATAGTTCAGAAGTA
GGAGCTGTTGCTTCTAATATCTCAGCAATAATCAACCATTTCAGCATCTGGCTTGCTACT
AGCCTCAGCATATTTTATTTGCTCAAGATTGCCAATTTCTCCAACCTTATTTTTCTCCAC
TTAAAGAAGAGAATTAGGAGTGTTGTTCTGGTGATACTGTTGGGTCCCTTGGTATTTTTG
ATTTGTAATCTTGCTGTGATAACCATGGATGACAGTGTGTGGACAAAAGAATATGAAGGA

AATGTGACTTGGA●ATCAAATTGAGGAATGCAATACACCTTTC●ATATGACTGTAAGC
ACACTAGCAAACCTCATACCCTTCATTCTGACCCTAATATGTTTTCTGCTGTTAATCTGT
TCTCTGTGTAAACATCTCAAGAAGATGCAGCTCCATGGCAAAGGATCTCAAGATCCCAGC
ACCAAGGTCCACATAAAAGCTTTGCAAACTGTGACCTCCTTTCTTCTGTTATGTGCCATT
TACTTTCTGTCCATGATCATATCAGTTTGTAATTTTGGGAGGCTGGAAAAGCAACCTGTC
TTCATGTTCTGCCAAGCTATTATATTCAGCTATCCTTCAACCCACCCATTCATCCTGATT
TTGGGAAACAAGAAGCTAAAGCAGATTTTTCTTTCAGTTTTGCGGCATGTGAGGTACTGG
GTGAAAGACAGAAGCCTTCGTCTCCATAGATTCACAAGAGGGGCATTGTGTGTCTTCTAG

SEQ ID NO:53

Human T2R31 amino acid sequence

MTTFIPIIFSSVVVVLFVIGNFANGFIALVNSIERVKRQKISFADQILTALAVSRVGLLW
VLLLNWYSTVFNPAFYSVEVRTTAYNVWAVTGHFSNWLATSLSIFYLLKIANFSNLIFLH
LKRRVKSVILVMLLGPLLFLACQLFVINMKEIVRTKEFEGNMTWKIKLKSAMYFSXMTVT
IGAXLVPFTLSLISFLMLICSLCKHLKKMQLHGEGSQDLSTKVHIKALQTLISFLLLCAI
FFLFLIVSVWSPRRLRNDPVVMVSKAVGNIYLAFDSFILIWRTKKLKHTFLLILCQIRC

SEQ ID NO:54

Human T2R31 nucleotide sequence

ATGACAACTTTTATACCCATCATTTTTTCCAGTGTGGTAGTGGTTCTATTTGTTATTGGA
AATTTTGCTAATGGCTTCATAGCATTGGTAAATTCCATTGAGCGGGTCAAGAGACAAAAG
ATCTCTTTTGCTGACCAGATTCTCACTGCTCTGGCGGTCTCCAGAGTTGGTTTGCTCTGG
GTATTATTATTAAATTGGTATTCAACTGTGTTTAATCCAGCTTTTTATAGTGTAGAAGTA
AGAACTACTGCTTATAATGTCTGGGCAGTAACCGGCCATTTCAGCAACTGGCTTGCTACT
AGCCTCAGCATATTTTATTTGCTCAAGATTGCCAATTTCTCCAACCTTATTTTTCTTCAC
TTAAAGAGGAGAGTTAAGAGTGTCATTCTGGTGATGCTGTTGGGGCCTTTACTATTTTTG
GCTTGTCAACTTTTTGTGATAAACATGAAAGAGATTGTACGGACAAAAGAATTTGAAGGA
AACATGACTTGGAAGATCAAATTGAAGAGTGCAATGTACTTTTCANATATGACTGTAACC
ATTGGAGCANACTTAGTACCCTTTACTCTGTCCCTGATATCTTTTCTGATGCTAATCTGT
TCTCTGTGTAAACATCTCAAGAAGATGCAGCTCCATGGAGAAGGATCGCAAGATCTCAGC

ACCAAGGTCCACA●AAAGCTTTGCAAACTCTGATCTCCTTCCT●TGTTATGTGCCATT
TTCTTTCTATTCCTAATCGTTTCGGTTTGGAGTCCTAGGAGGCTGCGGAATGACCCGGTT
GTCATGGTTAGCAAGGCTGTTGGAAACATATATCTTGCATTCGACTCATTCATCCTAATT
TGGAGAACCAAGAAGCTAAAACACACCTTTCTTTTGATTTTGTGTCAGATTAGGTGCTGA

SEQ ID NO:55

Human T2R32 amino acid sequence

HSFMLTMGSRKPKQTFLSAL

SEQ ID NO:56

Human T2R33 amino acid sequence

MVYFLPIIFSILVVFAFVLGNFSNGFIALVNVIDWVKRQKISSADQILTALVVSRVGLLW
VILLHWYANVFNSALYSLEVRIVASNISAVINHFSIWLAASLSIFYLLKIANFSNLIFLH
LKKRIKSVVLVILLGPLVFLICNLAVITMDERVWTKEYEGNVTWKIKLRNAIHLSSLTVT
TLANLIPFTLSLICFLLLICSLCKHLKKMQLHSKGSQDPSTKVHIKALQTVISFLMLCAI
YFLSIMISVWNLRSLENKPVFMFCKAIRFSYPSIHPFILIWGNKKLKQTFLSVFWQVRYW
VKGEKPSSP

SEQ ID NO:57

Human T2R33 nucleotide sequence

ATGGTATATTTTCTGCCCATCATTTTTTCCATTCTGGTAGTGTTTGCATTTGTTCTTGGA
AATTTTTCCAATGGCTTCATAGCTCTAGTAAATGTCATTGACTGGGTTAAGAGACAAAAG
ATCTCCTCAGCTGACCAAATTCTCACTGCTCTGGTGGTCTCCAGAGTTGGTTTACTCTGG
GTCATATTATTACATTGGTATGCAAATGTGTTTAATTCAGCTTTATATAGTTTAGAAGTA
AGAATTGTTGCTTCTAATATCTCAGCAGTAATCAACCATTTCAGCATCTGGCTTGCTGCT
AGCCTCAGCATATTTTATTTGCTCAAGATTGCCAATTTCTCCAACCTTATTTTTCTCCAC
CTAAAGAAGAGAATTAAGAGTGTTGTTCTGGTGATACTGTTGGGGCCCTTGGTATTTCTG
ATTTGTAATCTTGCTGTGATAACCATGGATGAGAGAGTGTGGACAAAAGAATATGAAGGA

AATGTGACTTGGA[redacted]ATCAAATTGAGGAATGCAATACACCTTTC[redacted]GCTTGACTGTAACT
ACTCTAGCAAACCTCATACCCTTTACTCTGAGCCTAATATGTTTTCTGCTGTTAATCTGT
TCTCTTTGTAAACATCTCAAGAAGATGCAGCTCCATAGCAAAGGATCTCAAGATCCCAGC
ACCAAGGTCCACATAAAAGCTTTGCAAACTGTGATCTCCTTCCTCATGTTATGTGCCATT
TACTTTCTGTCCATAATGATATCAGTTTGGAATCTTAGGAGTCTGGAAAACAAACCTGTC
TTCATGTTCTGCAAAGCTATTAGATTCAGCTATCCTTCAATCCACCCATTCATCCTGATT
TGGGGAAACAAGAAGCTAAAGCAGACTTTTCTTTCAGTTTTTTGGCAAGTGAGGTACTGG
GTGAAAGGAGAGAAGCCTTCATCTCCATAG

SEQ ID NO:58

Human T2R34 amino acid sequence

GSSRXKPPRIPHKKLCKLGPSFPHNNLPIYFLCXNHIVLEFLKMRPKKKCSLMLCQAFGI
IYPSFHSFILXWGNKTLKQTFLSVXWQVTCWAKGQNQSTP

SEQ ID NO:59

Human T2R35 amino acid sequence

NAIRPSKLWTVTEADKTSQPGTSANKIFSAGNLISHVNMSRRMQLHGKGSQHLSTRVHIK
AXQTVISFLMLXAIYFLCLITSTWNPRTQQSKLVFLLYQTLGFMYLLFHSFILTMGSRKP
KQTFLSAL

SEQ ID NO:60

Human T2R36 amino acid sequence

MICFLLIILSILVVFAFVLGNFSNGFIALVNVIDWVKRQKISSADQILTALVVSRVGLLW
VILLHWYSNVLNSALYSSEVIIFISNAWAIINHFSIWLATSLSIFYLLKIVNFSRLIFHH
LKRKAKSVVLVIVLGPLVFLVCHLVMKHTYINVWTKEYEGNVTWKIKLRNAIHLSNLTVS
TLANLIPFTLTLISFLLLIYSLCKHLKKMQLHGKGSQDPSTKVHIKALQTVTSFLLLCAI
YFLSMIISVCNFGRLEKQPVFMFCQAIIFSYPSTHPFILILGNKKLKQIFLSVFWQMRYW
VKGEKPSSP

SEQ ID NO:61

Human T2R36 nucleotide sequence

ATGATATGTTTTCTGCTCATCATTTTATCAATTCTGGTAGTGTTTGCATTTGTTCTTGGA
AATTTTTCCAATGGCTTCATAGCTCTAGTAAATGTCATTGACTGGGTCAAGAGACAAAAG
ATCTCCTCAGCTGACCAAATCCTCACTGCTCTGGTGGTCTCCAGAGTTGGTTTACTCTGG
GTAATATTATTACATTGGTATTCAAATGTGTTGAATTCAGCTTTATATAGTTCAGAAGTA
ATAATTTTTATTTCTAATGCCTGGGCAATAATCAACCATTTCAGCATCTGGCTTGCTACT
AGCCTCAGCATATTTTATTTGCTCAAGATCGTCAATTTCTCCAGACTTATTTTTCATCAC
TTAAAAAGGAAGGCTAAGAGTGTAGTTCTGGTGATAGTGTTGGGTCCCTTGGTATTTTTG
GTTTGTCACCTTGTGATGAAACACACGTATATAAATGTGTGGACAAAAGAATATGAAGGA
AATGTGACTTGGAAGATCAAACTGAGGAATGCAATACACCTTTCAAACTTGACTGTAAGC
ACACTAGCAAACTTGATACCCTTCACTCTGACCCTGATATCTTTTCTGCTGTTAATCTAC
TCTCTGTGTAAACATCTCAAGAAGATGCAGCTCCATGGCAAAGGATCTCAAGATCCCAGC
ACCAAGGTCCACATAAAAGCTTTGCAAACTGTGACCTCCTTTCTTCTGTTATGTGCCATT
TACTTTCTGTCCATGATCATATCAGTTTGTAATTTTGGGAGGCTGGAAAAGCAACCTGTC
TTCATGTTCTGCCAAGCTATTATATTCAGCTATCCTTCAACCCACCCATTCATCCTGATT
TTGGGAAACAAGAAGCTAAAGCAGATTTTTCTTTCAGTTTTTTGGCAAATGAGGTACTGG
GTGAAAGGAGAGAAGCCTTCATCTCCATAG

SEQ ID NO:62

Human T2R37 amino acid sequence

MITFLPIIFSILIVVTFVIGNFANGFIALVNSIEWVKRQKISSADQISHCSGGVQNWFTL
GHIITLVCNCV*FGFI*IRSKNFWF*CLSNNQAFQHVGVTSLSIFHLLKTANFSNLIFLH
LKKRIKSVGLVILLGPLLFFICNLFVINMDESVWTKEYEGNVTWKIKLRSAMYHSNMTLT
MLANFVPFTLTLISFLLLICSLCKHLKKMQLHGKGSQDPSTKVHIKALQTVTSFLLLCAI
YFLSMIISVCNLGRLEKQPVFMFCEAIIFSYPSTHPFILILGNKKLKQIFLSVLRHVRYW
VKGEKPSSS

SEQ ID NO:63

Human T2R37 nucleotide sequence

ATGATAACTTTTCTGCCCATCATTTTTTCCATTCTAATAGTGGTTACATTTGTGATTGGA
AATTTTGCTAATGGCTTCATAGCTCTAGTAAATTCCATTGAGTGGGTTAAGAGACAAAAG
ATCTCATCAGCTGACCAAATTTCTCACTGCTCTGGTGGTGTCCAGAATTGGTTTACTCTG
GGTCATATTATTACATTGGTATGCAACTGTGTTTAATTTGGCTTCATATAGATTAGAAGT
AAGAATTTTTGGTTCTAATGTCTCAGCAATAACCAAGCATTTCAGCATGTGGGTGTTACT
AGCCTCAGCATATTTCATTTGCTCAAGACTGCCAATTTCTCCAACCTTATTTTTCTCCAC
CTAAAGAAGAGGATTAAGAGTGTTGGTTTGGTGATACTATTGGGGCCTTTGCTATTTTTC
ATTTGTAATCTTTTTGTGATAAACATGGATGAGAGTGTATGGACAAAAGAATATGAAGGA
AACGTGACTTGGAAGATCAAATTGAGGAGTGCAATGTACCATTCAAATATGACTCTAACC
ATGCTAGCAAACTTTGTACCCTTCACTCTGACCCTGATATCTTTTCTGCTGTTAATCTGT
TCTCTGTGTAAACATCTCAAGAAGATGCAGCTCCATGGCAAAGGATCTCAAGATCCCAGC
ACCAAGGTCCACATAAAAGCTTTGCAAACTGTGACCTCCTTTCTTCTGTTATGTGCCATT
TACTTTCTGTCCATGATCATATCAGTTTGTAATTTGGGGAGGCTGGAAAAGCAACCTGTC
TTCATGTTCTGCGAAGCTATTATATTCAGCTATCCTTCAACCCACCCATTCATCCTGATT
TTGGGAAACAAGAAGCTAAAGCAGATTTTTCTTTCAGTTTTGCGGCATGTGAGGTACTGG
GTGAAAGGAGAGAAGCCTTCATCTTCATAG

SEQ ID NO:64

Human T2R38 amino acid sequence

MLTLTRIRTVSYEVRSTFLFISVLEFAVGFLTNAFVFLVNFWDVVKRQPLSNSDCVLLCL
SISRLFLHGLLFLSAIQLTHFQKLSEPLNHSYQAIIMLWMIANQANLWLAACLSLLYCSK
LIRFSHTFLICLASWSPGRSPVPS

SEQ ID NO:65

Human T2R39 amino acid sequence

LRNAGLNDSNAKL[redacted]NNDLLLINLILLLPLSVFVMCTSMLFVSL[redacted]MHWMQSESHKLSS
ARTEAHINALKTVTTFFCFFVSYFAAFMANMTFRIPYRSHQFFVVKEIMAAYPAGHSVII
VLSNSKFKDLFRRMICLQKE

SEQ ID NO:66

Human T2R40 amino acid sequence

SQYSLGHSYVVIFGYGQMKKTFLGILWHLKCGLKGRALLATQVGLREKSTRSLGVIFLAS
SYSFFVYVLCH

SEQ ID NO:67

Human T2R41 amino acid sequence

MITFLLIILSILVVFAFVLGNFSNGFIALVNVIDWVNTRKISSADQILTALAVSRVGLLW
VILLHWYANVLNPALYSSEVIIFISNISAIINHFSIWLATSLSIFYLLKIVNFSRLIFHH
LKRKAKSVVLVIVLGPLVFLVCHLVMKHTYINVWTKEYEGNVTWKIKLRNAIHLSNLTVS
TLANLIPFTLTLISFLLLICSLCKHLKKMQLHSKGSQDPSTKVHIKALQTVTSFLMLFAI
YFLYLITSTWNL*TQQSKLVFMFCQTLGIMYPSFHSFILIMGSRKLKQTFLSVLCQVTCL
VKGQQPSTP

SEQ ID NO:68

Human T2R42 amino acid sequence

FIGLTDCIAWMRNQKLCMVGFILTRMALARINIL

SEQ ID NO:69

Human T2R43 amino acid sequence

LELIFS*KVVATRGLVLGMLGNGLIGLVNCIEWAKSWKVSSADFILTSLAIVRIIRLYLI
LFDSFIMVLSPHLYTXXXXXXXXXXXXXXXXXXXXXXXXSLSIFHWFKTANFSNLIFLPLK

EED*NVWLGDAV[redacted]LGIFHL*SCSENHG*EVCGQKNMKEFCSG[redacted]LRNAIQLSNLTVTM
PANVTPCTLTLISFLLLIYSPCKHVKKMQLHGKGSQHLSTKVHIKVLQTVISFFLLCAIY
FVSVIISVWSFKNLENKPVFMFCQAIGFSCSSAHPFILTMGNKKLKQTYLSVLWQMR

5

SEQ ID NO:70

Human T2R44 amino acid sequence

MITFLPIIFSILIVVIFVIGNFANGFIALVNSIEWVKRQKISFVDQILTALAVSRVGLLW
10 VLLLHWYATQLNPAFYSVEVRITAYNVWAVTNHFSSWLATSLSMFYLLRIANFSNLIFLR
IKRRVKSVVLVILLGPLLFLVCHLFVINMDETVWTKEYEGNVTWKIKLRSAMYHSNMTLT
MLANFVPLTLTLISFLLLICSLCKHLKKMQLHGKGSQDPSTKVHIKALQTVTSFLLLCAI
YFLSMIISVCNLGRLEKQPVFMFCQAIIFSYPSTHPFILILGNKKLKQIFLSVLRHVRYW
VKDRSLRLHRFTRGALCVF

15

SEQ ID NO:71

Human T2R45 amino acid sequence

20 MATELDKIFLILAIAEFIISMLGNVFIGLVNCSEGIKNQKVFSADFILTCLAISTIGQLL
VILFDSFLVGLASHLYTTYRLGKTVIMLWHMTNHLTTWLATCLSIFYFFKIAHFPHSLFL
WLRWRMNGMIVMLLILSLFLLIFDSLVLEIFIDISLNIIDKSNLTLYLDESKTLYDKLSI
LKTLLSLTSFIPFSLFLTSLLFLFLSLVRHTRNLKLSSLGSRDSSTEAHRRAMKMVMSFL
FLFIVHFFSLQVANWIFFMLWNNKCIKFVMLALNAFPSCHSFILILGNSKLQQTAVRLLW
25 HLRNYTKTPNPLPL

SEQ ID NO:72

Human T2R46 amino acid sequence

30

MSFLHIVFSILVVVAFILGNFANGFIALINFIAWVKKQKISSADQIIADKQSPELVCSG

SEQ ID NO:73

Human T2R47 amino acid sequence

MLNALYSILIIIINI*FLIGILGNGFITLVNGIDWVKM*KRSSILTALTISRICLISVIM
VRWFI

SEQ ID NO:74

Human T2R48 amino acid sequence

VSRVGLLWVILLHWYSTVLNPTSSNLKVIIFISNAWAVTNHFSIWLATSLSIFYLLKIVN

SEQ ID NO:75

Human T2R49 amino acid sequence

TVTMLANLVPFTVTLISFLLLVCSLCKHLKKMHLHGKGSQDPSTKVHIKVLQTVISFLLL
CAIYFVSVIISS

SEQ ID NO:76

Human T2R50 amino acid sequence

MITFLPIIFSILVVVTFVIGNFANGFIALVNSTEWVKRQKISFADQIVTALAVSRVGLLW
VLLLNWYSTVLNPAFYSVELRTTAYNIWAVTGHFSNWPATSLSIFYLLKIANFSNLIFLR
LKRRVKSVILVVLLGPLLFLACHLFVVNMNQIVWTKEYEGNMTWKIKLRRAMYLSDTTVT
MLANLVPFTVTLISFLLLVCSLCKHLKKMQLHGKGSQDPSTKVHIKVLQTVISFFLLCAI
YFVSVIISVWSFKNLENKPVFMFCQAIGFSCSSAHPFILIWGNKKLKQTYLSVLWQMRY

SEQ ID NO:77

Rat T2R01 amino acid sequence

MMEGHILFFFLVVMVQFVTGVLANGLIVVVHAIDLIMWKKMAPLDLLLFCLATSRIILQL
CILFAQLCLFSLVRHTLFEDNITFVFIINELSLWFATWLGVFYCAKIATIPHPLFLWLKM

RISRLVPWLILG[redacted]LYVIITTFIHSRETSAILKPIFISLFPKNA[redacted]VGTGHATLLSVLVL
GLTLPLFIFTVAVLLLIYSLWNYSRQMRTMVGTREYSGHAHISAMLSILSFLILYLSHYM
VAVLISTQVLYLGSRTFVFCLLVIGMYPSIHSIVLILGNPKLKRNAKMFIVHCKCCHCTR
AWVTSRSPRLSDLPVPPTHPSANKTSCSEACIMPS

SEQ ID NO:78

Rat T2R01 nucleotide sequence

CAGGAATCATAAATGGCTGAAACTGGGCAGAACTCTATGCATTATTTAAAGAAGTCATTG
GTTTGTCATTCTTAAA**ATGATGGAAGGGCATATACTCTTCTTCTTTTTGGTTGTGATGGT**
GCAGTTTGTCACTGGGGTCTTGGCAAATGGCCTCATTGTGGTTGTCCATGCTATTGACTT
GATCATGTGGAAGAAAATGGCCCCGTTGGATCTGCTTCTATTTTGCCTGGCGACTTCTCG
GATCATTCTGCAGTTATGTATATTGTTTGCACAATTGTGTCTATTCTCTTTGGTGAGACA
CACTTTATTTGAGGACAATATTACCTTTGTCTTCATCATAAATGAACTGAGTCTTTGGTT
TGCTACATGGCTCGGTGTTTTCTACTGTGCCAAGATTGCTACCATTCCTCACCCACTCTT
TCTGTGGCTGAAGATGAGGATATCCAGGTTGGTACCATGGCTGATCCTGGGATCTGTGCT
CTATGTAATTATTACTACTTTCATCCATAGCAGAGAGACTTCAGCAATCCTTAAACCAAT
TTTTATAAGCCTTTTTCCTAAAAATGCAACTCAAGTCGGAACAGGGCATGCCACACTACT
CTCAGTCCTGGTCCTTGGGCTCACACTGCCGTTGTTCATCTTTACTGTTGCTGTTCTGCT
CTTGATATACTCCCTGTGGAATTATAGCAGGCAGATGAGGACTATGGTAGGCACCAGGGA
GTATAGCGGACATGCTCACATCAGTGCAATGCTGTCCATTCTATCATTCCTCATCCTCTA
TCTCTCCCACTACATGGTGGCTGTTCTGATCTCTACTCAAGTCCTCTACCTTGGAAGCAG
AACCTTTGTATTCTGCTTACTGGTTATTGGTATGTACCCCTCAATACACTCGATTGTCTT
AATTTTAGGAAATCCTAAGCTGAAACGAAATGCAAAAATGTTCATTGTCCATTGTAAGTG
TTGTCATTGTACAAGAGCTTGGGTCACCTCAAGGAGCCCAAGACTCAGTGACTTGCCAGT
GCCTCCTACTCATCCCTCAGCCAACAAGACATCCTGCTCAGAAGCCTGTATAATGCCATC
CTAATTGTCCAGCCTGAGGTTTAATCCTAGGTTTGGTACTATTTCAAAGAGTAAAGTTGA
TCATTAAAGCACAACATATGTTGGTGGATGACATCAAGGTCCATATCCCAGTTGTCAATT
GTAAACCTCACCTTGCAAGATGATGTCACTGAGAAAGCAGGACAAATGGAGTCTAGGTCC
TTCTGTATGACTTGCTGCAGTATATGTGAATCTATAATTTTCTCCAAAAAAACAAAAAAA
AAAAAAAAAAA

SEQ ID NO:79

Rat T2R02 amino acid sequence

MFSQKTNYSHLFTFSIIFYVEIVTGILGNGFIALVNIMDWLKRRRISTADQILTALALTR
LIYVWSVLICILLLFLCPHLSMRPEMFTAIGVIWVVDNHFSIWLATCLGVFYFLKIASFS
NSLFLYLKWRVKKVVLMIILISLIFLMLNISSLGMYDHFSIDVYEGNMSYNLVDSTHFPR
IFLFTNSSKVFLIANSSHVFLPINSLFMLIPFTVSLVAFFVLFLSLWKHHKKMQVNAKGP
RDASTMAHTKALQIGFSFLLLYAIYLLFIITGILNLDLMRCIVILLFDHISGAVFSISHS
FVLILGNSKLRQATLSVLPCLRCRSKDMDTVVF

SEQ ID NO:80

Rat T2R02 nucleotide sequence

ATTTTGCTCCACTATTTTGCTCTTCTGCAGTAACACAGACCACAAAACAATGGAGCCAAT
GGGTCAAGAGCTGAAACTTCAGGAAGTGGGAGCCAAATTTTCTTTGTGATAGGTTGGCAT
ATGAGAATTCATTATTTGATGCAGCTTCTGAAAACTGGATGTGAAATACTGGATGAAGCA
GAGGTGATGACCCCTTTGAAATTAAAAAGCCAAGATGTTCATGGAGAAATTATAAAACAA
TATCTGGGAAATTTGATGCTTCCTAATCGGGTGTAAATGGGATTTTAAATGATGAACATT
TTGAATTTCCAATGACCATTATGTAAAGTTTTTAAACACAGTAGAGACATCATAAATTGA
AGC**ATGTTCTCACAGAAAACAAACTACAGCCATTTGTTTACTTTTTCAATTATTTTTTAT**
GTGGAAATAGTAACAGGAATCTTAGGAAATGGATTCATAGCACTAGTGAATATCATGGAC
TGGCTCAAGAGGAGGAGGATCTCTACTGCAGATCAGATTCTCACTGCTTTGGCCCTTACC
AGACTCATTTATGTGTGGTCTGTACTCATTTGTATATTGTTACTATTTCTGTGCCCACAT
TTGTCTATGAGACCAGAAATGTTTACAGCGATAGGTGTTATCTGGGTAGTGGATAACCAC
TTCAGCATCTGGCTTGCTACATGTCTTGGTGTCTTTTATTTCCTCAAAATAGCCAGTTTT
TCTAACTCTTTGTTTCTTTACCTAAAGTGGAGAGTTAAAAAAGTGGTTTTAATGATAATA
CTGATATCACTGATTTTCTTGATGTTAAACATTTCATCATTAGGGATGTATGATCATTTC
TCAATTGATGTTTATGAAGGTAATATGTCTTATAATTTGGTGGATTCAACACATTTTCCC
AGAATTTTCTTATTCACAAACTCATCTAAGGTCTTCTTAATCGCCAATTCATCCCATGTT
TTCTTACCCATCAACTCACTCTTCATGCTCATACCCTTCACAGTTTCCCTGGTAGCTTTT
TTCGTGCTCTTTCTCTCACTGTGGAAGCATCACAAGAAGATGCAGGTCAATGCCAAAGGA
CCCAGAGATGCCAGCACCATGGCCCACACAAAAGCCTTGCAAATTGGGTTCTCCTTCCTC
CTGCTGTATGCAATATACTTACTTTTCATTATCACAGGAATTTTGAACCTTGACTTGATG

AGATGTATAGTAA●CTTTTATTTGACCACATATCTGGAGCAGT●TTTCTATAAGCCAC
TCATTTGTGCTGATTCTGGGAAACAGTAAGCTGAGACAAGCCACTCTTTCTGTGCTGCCT
TGTCTTAGGTGCCGGTCCAAAGATATGGACACTGTCGTTTTCTAATAAATTCCAGAGTAC
ATTATGCAAAATCTTGAGGGTGATCAGTTCATAGAAAAAGTAATCTTAGAGGGGAAAATA
AAATATTGGGGCTTCAAATGTTGGATGGGTAATACATAGGAAGGCAGGACAAGGATGAAG
GAGACTAGCATTATATAAGTGATTTCACAGGGGAAATGGGAAAGAGGGCTTTTATATAAT
GAAGAAGAAGATAAATGATGAAGGATGAGGAAGAGTTAAATATGTAAAATGACAATAGAG
ATGGCATCATGCCGTTTTAAGAAATTTGGAATGCATATGTATGTTTATATATTTTTTAAT
TTTTATTGAATATATTTATTTACATTTTAAATGTTATCCTGTTTCCCCCACCCAACCTCC
CACCTCTTCCCACCTCCTTGCCCTGACATTCCCCTGCACTGGGGAATCCAGCCTTGACAG
GACCAAGGGCTTCTCCTCCCTTTGTTGCCAACAAGGCCATTCTTTGCTACATGTGCAGCA
GGAGCCATGGATCTGTCTATGTGTACTCTTTGGATGGTGGTTTAGTCCCTGGGAGCTCTT
GTTGGTTGGTATTGTTGTTCTTATGGTGTTGCAACTCCCTTCAGCTCCTTCAATCCTTCC
TGTAACTCCTCCAATGTGGACCCTGTTCTCAGTCCAATGGTTGACTATGAGCATTCACCT
CTGTGATTGTCATGCTCTGGCACAGCTTCTCAGAAGACAGCTACATCAGTCTCCTATAAG
AGTGCACTTCATGGCATCAGCAATGTTGTCTTGATTTGGTGTCTGTATGTATATGGGCTG
GATCCCAGGTGGGGCAGGCGCTGAATGGTCATTCCTTCAGTCTTTGCTCCAAACTTTGTC
TTTATATCTCCTATGAATATTTTTGTTCCCCCTTATAAGAATGACTGAAGTATCCACACT
TTGGCCATCCTTCTTCATGAGCTTCATGTGGTCTGTGAATTGTACATTGTGTAATCCAAG
CTTTTGGGCTAATATCCAATTATAGTGAGTGCATACCAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

SEQ ID NO:81

Rat T2R03 amino acid sequence

MVPTQVTIFSIIMYVLESLVIIVQSCTTVAVLFREWMHFQRLSPVEIILISLGISHFCLQ
WTSMLYNFGTYSRPVLLFWKVSVVWEFMNVLTFWLTSLLAVLYCVKVSSFSHPVFLWLRL
KILKLVLWLLLGALIASCLSIIPSVVKYHIQMELLTLDHLPKNSSLILRLQMFEWYFSNP
FKMIGFGVPFLVFLISIILLTVSLVQHWGQMKHYSSSSSSLRAQCTVLKSLATFFIFFTS
YFLTIVVSFIGTVFDKKSWFWVCEAVIYGLVCIHFTSLMMSNPTLKKALRLQFWSPESS

SEQ ID NO:82

Rat T2R03 nucleotide sequence

GCATGGTGCCAACCCAAGTCACCATCTTCTCTATCATCATGTATGTGCTTGAGTCCTTAG
TCATAATTGTGCAAAGTTGCACAACGGTTGCAGTGCTGTTCAGAGAGTGGATGCACTTTC
AAAGACTGTCGCCGGTGGAAATAATTCTCATCAGCCTGGGCATTTCACATTTCTGTCTAC
AGTGGACATCGATGCTGTACAACTTTGGTACCTACTCTAGGCCTGTCCTTTTATTTTGGA
AGGTATCGGTCGTCTGGGAGTTCATGAACGTTTTGACATTCTGGCTAACCAGTTTGCTTG
CTGTCCTCTACTGTGTCAAGGTCTCTTCCTTCTCTCACCCCGTCTTCCTCTGGCTGAGGT
TGAAAATTTTGAAACTGGTTCTCTGGTTGCTATTGGGCGCTCTGATAGCTTCTTGTTTGT
CAATCATCCCTTCTGTTGTTAAATATCATATCCAGATGGAATTACTCACCCTAGATCATT
TACCCAAAAACAGTTCTTTGATTCTAAGACTGCAAATGTTCGAGTGGTATTTTTCTAATC
CTTTCAAAATGATTGGGTTTGGCGTTCCTTTCCTCGTGTTCCTGATTTCTATCATCTTAC
TCACAGTCTCGCTGGTCCAGCATTGGGGGCAGATGAAACACTACAGCAGCAGCAGCTCCA
GCCTGAGAGCTCAGTGCACTGTTCTGAAGTCTCTTGCCACCTTCTTCATCTTCTTCACAT
CCTATTTTCTGACTATAGTCGTCTCCTTTATTGGCACCGTGTTTGATAAGAAGTCATGGT
TCTGGGTCTGCGAAGCTGTCATCTATGGTTTAGTCTGTATTCACTTCACTTCCCTGATGA
TGAGCAACCCTACACTGAAAAAAGCACTCAGGTTGCAGTTCTGGAGCCCAGAGTCTTCCT
AAGGCAGGGAATTCAGTGAAGCCTCTGGGGTAAGGAGGCTTTGCATTGGCACAGTTCTTA
GAGTGAAATGCAAACGTGGACACGAACTTCATTCTCTTTCATGTCCACAGATGGATGGAT
CTATAAATCATCACCAATCTTCCCTGTATTCTGACCCATCCTTTTCCTGTCCTATCCATA
GTCCCCAGGTTGGTTTTGATTTTTCTCATGATCACACCTTAGCTTTAGCCACCGTTGCAA
TATCAAACATGATCTATATGTTACAGCCAAAATCATTCTCACAATTGTCAATTGCTTCAC
AAATTCAGATAAATCCCCCTTCCTGTCAGGAATGTATTGTCTGTGCATTCAATGCTCACC
ATGCTAAGCCATTCATTCCCTTCCTAACTTGAGTTTAAGAAGAAAATGTCTTACTGTTGC
CCATGTCCTATTGTGCTGCTTCTGGATGTTTTATGCAGTGATTTAGACACACGCCCTTGC
CTGTCTCCAAATACTGGCCCTTTATTCCTTTATAAGTCTAGTAGAAAATGAACTCGTCTT
TACTTCATTGACGAAGACATTGTATTCTTCCCCAAAATAGTGTTTAACTACTCTAGTCTC
ATCCATAATATCCCTAAATATCAGTGATTTCAGTGAGTAAAACCTGACAACAGTTATTGC
TTTGACTCTTAATTCAATTGTGCTGTAACATAGAGGAAACATTCTAGAACATTTCCATAT
TAATTTGTGCTTGTAGCAAACCAAAATTCTCCCCAGTTGGGTAAAAATATCAAAAGCACA
GAGTAATCAATTTTGAAATCACTCAGAAGACATCATTGTTCTATATATGTTTTTTTTAAA
CTTCCCTCTAACAAGTATCAGATCTTTGCCTTTACAGGGTCTGGTCTTACCATGACTATA
TTTTATCACCATGACCTATTTTCTCTTCATCTCTTTGTTTTCACTAACTCAGTAGCAACC
AAATATCACATTAATAGCTAACTCTGGGCACTTATTTCTCAGCCTTTATCTATTCCAGAC

ACTTTCAATGTA[redacted]TCTGCTAAACACAATGACATCTCTTTTTGT[redacted]TTCTAACGACAAGGA
ATCATAACTTTCCAACTTTTATACATGGTAGACATATTTGGTGAACTTAACTTCTGACTC
TTTCTTTAGAAGACTGAAACTACTCCGGAAAGCAAGCCTTCTGATGGAGAAATAGATACG
GGTATCGTGATTCATTGTGAAAGTGAATTCCGGTGCCTGGAAAGAAATGGATATTTTTTT
TTCTCTTGAGTGTGTCACTCTGACATATGTTCCATGTTGAATCCATATTTGATACTGATA
GCATGAATGTAAGTAAAGCATGTATGTAAGTAAAGACTGCTACCAAAACTTCGATTCAAC
TTTCCTCAGCAGTATCCCTGATATTGCATAAGAAAGAAAAAACACGCTGTCCTACTTGAA
GAAGGACGTGTTCCATGCAATGTGGATGTGTCCCAGGCTACATTGGCTCAACTGCAGCTG
AAGGTGGGATGGGAAATGGTATAGTTAGTAATGTCTGCTGAGCTGTCTCACTGGAAAGGA
TTCTGAGCAGAGTAAATGTAAGCAATGTGGCCAAGGTCTCCTAGGAATGGGTTGTAAGCT
TGTAAGGAGTTGGGTTGTAAGAGTTTGGGATCCTTTCAGAATGGATTGAGCAAGAGCCAC
TGAAACTTGGACTATACCTTTGTTATTTGTATCTAAATCCAGAAGGGTCTTTGCATGTTC
CAAAATCTCAGATAGCTGGAAGGAAGAAGGACTGTTCTCTTTACAAGTATATAAATAGAG
AATGAGCTAAAAAGGACCCCCTCACCCCCGCCGTCACACACAGGAATACTATTCCAGAAA
CTAGGGAGTATTTTTAGTGTTCTCACTATTTCCCTTTGAAAAAAGTGCAATGGAAAACTT
ATCCATGACATACATGAGGTTGGAGTGATAAAAACAGCTGAAGGAAGAGGAAGTCTGAAA
AAAGATGGAAACAGCAATGATGCTTGTCCTATATATGTGTGACACCCACTAGTTCCCAAG
GAAACCTTACATCCATTATCTCATTTCAAGCTGGAAGGACAAGTCAAGATCACTCAACCG
ACCCAGCTGGAAAACAGACCTAAGAATGTTAAACTCATACTGATGGTTATTTCTCACTCT
AAAGTCAATGCAAATGGATAGCAAACAAAGGGGCTATTTTTTTAAGGGACCAGAGGGTTT
CAATCTAGAATCAGAGAAAAGATAAAAAGGGAGATGCTATAGAAAAACAATAGAGAAGAT
GTGGCCAAGAACAAGGAAAATCTCCAGTTAGCTTGGCACTTAGGGGCCAACATGTTTCTG
TTGTTCGGTCTTCAATACTGTATTGCATGTTGGGCTCACTATGTTTTAGTTGTGAGTGGG
TTGTGCTTCCTGGAATTAAGAAAGGTCTGTTTCTAGATTTCAGGTACAAATGTTTAGAAG
CCCATTGGTAGCATCAGTGAAATTAGGAAAAAACTGTGAGCACTGCTGGCTGGACTTGGC
AAAGTCATTCACTATTTACACATCAAATTATTAGCAACTTGAAAGTAAATCTTTGCTCAT
CATCCAGTGGCCCCCATGATCCTGGTGAATGACTTGTAATACTGTGGAGACTGGCAACGA
CGGTGAATTCCTAGTAACACTTACCATAGAATCTGTTCATAATTAGACTCGCCCAGATTT
TAGTTGCTAGAGAACAATCTTTCTCCTTTACCCACATTCCTACTGAGTAGGATGCATAGG
TTCGGAAACCCCCATGGCATCGTTTGACTCCTCCTGGTAGTCAAGAGAGTCCAGTCACCA
GTCTCCGAAACACCTGCCAAGTCCTAACTCCCAACAGTCTACAGTGTAAACCTCAGTGTT
TGCATGAGGTTTATGTATCTCCTTACCATTTCCTAAATGTCAATACCCGTGCACAGGATA
TTTGCATAGGCTGCCTCCAAGCCTGGGAAACACTCTCCTCCTCGCATTTGCTGGGTTTCA
CCTTTCCAATTCAGTGTGCCCTTTAAAAGGCACTGCTTTTCTAGGCCCACCACTATTGCT

GCTCACGCATGA[redacted]ATCAAATCTACCACAGGCTTTTGCCTCTCA[redacted]ATTATTCTTCTTTC
TACTATGCAATGTGGTATCCATGAGAACTTTGTCACATTGTCAAATTCTACCTTTGTTTT
AATGnGnGCCTTTGTAATAGnGACTATGCCCAGAAATTAAATTATAGTAAGATGGGTAAC
AACnCTTCAATTnTGGAATTTATAATTAAATAAATATTATGTAATATTATGACTTATTAT
AAnGTCAATCTACTGTACCCTACTCCTACTAGGAATGCAAAGACAAATAGCAATGTGATC
AGCATGTGCTCTTTCACAAGATCATATTGTGCATGTTGCTGATGATGCCCACAGTGCATC
TATCAGAATATCTCTGATCATTTTTTTTTTTTGCTTTTGAGAAGCCCCGGTTGGTGCTG
GGATGCTTCATAGCAGGTCCACCATAGACACATGCTTAGAGGAAAGCTGCCTCTCTCTCT
TCATTCCCAAGGAACAGTAAAAGCAGAAAAGGCTCTTATGTTCTAAAGAACAGAAAATAG
CCTGCATTTCAACTACCTCCTGTTCAGAAGGCACCGAAACACACCACCAAGCAAGACACC
CCTTTACTTTCTCCTGCTTCCCTCAATTTGATGATCATTTGGAAATAAGAAGAAAGAAAA
AGATGTGGAAGCCAATTAAAAACAGTCTTGTCTATCTCCCTGGTGAGCTCTCAACTTCTT
AGTCAGACCAAAGTAGGTGAAAAAATAATAATTTTTAATTTGGTATGAGAGTCATGTTTA
GGCTGAAAATCTTAAAAAATCTTAGCATAAAAACATTTTCCCCTAGACCCATGAAATTTA
TAATATTATCTGTGGTTGAGAAAGGCTAGTTATAGAAAAATGTTTAGAATCAGAATATTT
TGAGGGCTCTTTTTTTGTTTTGCCTAATCATTACATTTGTTATAAGAAGTCTAAAAGTTG
GTATGCTACAGGTCTTGTCATATTTTCTCTGAGGTTGAGTGCCAAGTAGTCTGCATTGTG
TTTAAATCCTGCTTAAAATTATCCCAAGACAATATAACTTCTCAGGAGCTAAGCCAAGGG
CCCCTTTCAGACTACCTTAGTCCTCTCTCACCGTTGTCACCGTGGCTCATACATCAGAAT
CCTGAGGGAGCATCATGAAATCTAAGGCTTTACAACAGAATCTTTCTATCCCTGGTAGAA
ATCTTTTAACCTTGGGTTTTATTCTCATGCCATTCTGATGCTCGTATTTAAATTTTATGT
GTTTTTTCATATGTTCTTGCATTTCTATCGTTAAATTATGGTGACATACTTTCAAATGCT
TTGTTATTTTAAAAAGGGACAAAGAGAGATAGAAAGACAGGGAAAGATAGACAGAGGCTT
GCCTAATACAGTCAAGAAAGAAGCTATCAAAAGTATTTAGCAATACAACATTTATGATAT
ATTCATAACTGTTAACCATTTTTAATATTCTAAAATTTCACTTTTGTTTCAGAAATGTAT
ATTAAGAGAATCTGAGAAACATTTTTTTCTCATAGATGTAGAAAAACACACAAAATAAGG
TATAACACATTTAAGTGATTGAAAATAAAAACAAAAGCTTGCAAACAGGAGGAAAAGTAC
ATTGTAGGCTTTCGACATGGAGCTGCTACTAGGACCCAGGACTTGTTTATCATTTATTTG
CCAAGTCCCACAAACTCAGGGCAATACATCTCTGAGACAGTTTCCTATATTTTAATAAAA
CTTCCAAAATTGATACTCAGTGTGAATTGGCTAGCTTTAATGGCAGTCATTGGATAAACA
ATTCCAATGCCAAATTTCCCTAAGTTGATATATTTGATTAATATGTATATTAAAACATCA
GGCTATCCATCGGTTGGATCAAATACATTCTTTAGGGATCCATTCTTTTCCTTAAATTTG
ACTTATATGTGGATTCTTTTCACAATAAATAAGTAAATGAGCATTTATTTTAAAACTATT
TTAGACGGAACTGAATTACAGCCAAGGTAGTCAAAATGACTGAGAATAATCACTTACATA

TTTACAAGGGAAA[redacted]TGACTCTTCAGATTTAAGTTTAAAATTAGA[redacted]AGAGATAAATTTCA
CAAGCTTTCACTCCTAAGGCTAAAGATAGGCTGTGTAGGTAGTTATTTCTGAGCACATTG
GCACATCACCATTGTCAGTACTTGAGGGTTTGAATGAAGCTCACTCAAAGAACTTGGAAA
GAAGGTGGTCTTCTGACATCAATCAAGAAACAAGCTTTCCTCCCTACTTCTTCCCTAAAT
GCAACAACCTAAGAATTATCCACAAGATGGATGGCGCAAGGGTTCCTCAATCAATTTCAG
GATGTACATCAATGCGCAGCCTATACTACACCGAAAAGGAAGCGCATGGGTCTTAAAAAG
TAAAGGGGATATCAAAAAATTCGCAACCAAACAAAAAGTGGCACACATTTAAGCTAGGTC
TATGTTTGGTCAGTTACACCTGGAGAAGGGGGACATTTGGTCAGCTCATTCGAACACTGT
CAAGTCCTACCAACAATTCCTCTATGCTATTACCCATTAAACCTCAGGTCTCATCGAAAA
AAAAAAAAAAAA

SEQ ID NO:83

Rat T2R04 amino acid sequence

MLSAAEGILLCVVTSEAVLGVLGDTFIALANCMEYAKNKKLSKIGFILIGLAISRIGVVW
IIILQGYMQVFFPHILTFGNITEYITYIWVFLNHLSVWFATNLNILYFLKIANFSNSVFL
WLKSRVRVVFIFLSGCLLTSWLLCFPQFSKMLNNSKMYWGNTSWLQQQKNVFLINQSLTN
LGIFFFIIVSLITCFLLIVFLWRHIRQMHSDGSGLRDLNTEAHVKAMRVLISFAVLFILH
FVGLSIQVLCFFLPQNNLLFITGLIATCLYPCGHSIILILGNKQLKQASLKALQHLTCCE
TKRNLSVT

SEQ ID NO:84

Rat T2R04 nucleotide sequence

TGGTTCCATCACATGACAATAGGCTTGAAAAACTTGCAGATAGAGAAGACATAACCCCTC
CAACAAGAAGCCAACATATGGGACATTCTCCAGCAGATAATTTATAACAGATGCAACGGG
AGCAACTTCGAGATCTGCAAAG**ATGCTGAGTGCAGCAGAAGGCATCCTCCTTTGTGTTGT**
CACTAGTGAGGCAGTGCTGGGGGTTTTAGGAGACACATTCATTGCACTTGCAAACTGCAT
GGAGTATGCCAAGAACAAGAAGCTCTCTAAGATTGGTTTCATTCTCATTGGCTTGGCGAT
TTCCAGAATTGGTGTCGTATGGATAATAATTTTACAGGGGTATATGCAAGTATTTTTTCC
ACACATACTTACCTTTGGAAACATAACTGAATATATTACTTACATATGGGTGTTTCTCAA
TCACTTAAGTGTCTGGTTTGCTACCAACCTCAATATCCTCTACTTTCTAAAGATAGCAAA

TTTTTCCAACTCT[redacted]ATTTCTCTGGCTGAAAAGTAGAGTCCGTG[redacted]TTTTTATCTTTCT
GTCAGGATGCTTACTTACCTCGTGGTTACTATGTTTTCCACAATTTTCAAAGATGCTTAA
CAACAGTAAAATGTACTGGGGAAACACGTCTTGGCTCCAGCAGCAGAAAAATGTCTTCCT
TATTAACCAAAGTTTAACCAATCTGGGAATCTTCTTTTTCATTATTGTATCCCTGATTAC
CTGCTTCCTGTTGATTGTTTTCCTCTGGAGACACATCAGGCAAATGCACTCAGATGGTTC
AGGACTCAGAGACCTCAACACAGAAGCTCATGTGAAAGCCATGAGAGTTCTAATATCTTT
TGCGGTACTCTTTATCCTGCATTTCGTAGGTCTTTCCATACAAGTGCTATGCTTTTTTCT
GCCACAAAACAACCTACTCTTTATAACTGGTTTGATAGCCACATGCCTCTATCCCTGTGG
TCACTCAATCATCTTAATTCTAGGAAACAAGCAGCTGAAGCAAGCCTCCTTGAAGGCACT
GCAGCACTTAACGTGCTGTGAGACAAAAAGAAATCTCTCAGTCACATAAATGGGTTTGCC
AATTAATATCTGCCATGTTATTCCACTGATTTTTACCTGTTAGTTTCTCTGTGTCTCTGT
TTAGTTTCTGTTTCCATGATCTGTCCATTGATGAGCGTGGGGTGTTGAAATCTCCGACTA
TTGTTGTGTGAGATGAAATGTGTGCTTTGAGCTTTAGTAAGATTTCTTTTGTGAATGTAG
GTGCTTTTGCATTTGGTGCATAGATATTTAAGATTGAGAGTTCAGCTTGGTGGATTTTTC
CTTTGATGAATATGAAGTGTCCTTGCTTATCTTTTTTGATGACTTTTGATTGAACGTCAA
TTTTATTGGATATTAGATTGGCAACTCAAGATTGCTTCTTGAGGTCATTTGCTTGGAAAG
TTGTTTTTCAGCCATTTACTCTGAGGTAGTGTCTGTCTTTGTCTCTGAGGTGTGTTTCCT
GCATTCAGCAAAATGCTGGGTCCTCTTTACATATCCAGTTTGTTAGTCTATGTCTTTTTA
TTGGGGAATTGAGTCCATTGATGTTGAGAGATATTAATGAATAGTGATCATTGCTTCCTG
TTATTTTCGTTGTTAGATGTGGAATTATGTTTGTTTGTCTCTCTTTTGGTTTTATTGCAA
GGAAATTATATACTTGCTTTCTGTATGGTGTAGTTTCTCTCCTTGTGTTGCAGTTTTCCT
TCTATTATCCTTTGTAGGGCTAGATTTGAAGAAAGATATTGCATAAGCTTGGTTTTGTCA
TGGGATATCTTGGTTTCTCCATCTATGTTAATTGAGAGTTTTGCAGGATATAGTAGCCTG
GGATGACATTTGTGTTCTCTTAGGGTCTGTATGACATCTGTCCAAAATCTTCTGGCTTTC
ATAGTCTCTGGTGAGAAATCGGATGTAATTCTCATAAGTCTGCCATTATATGTCACTTGA
CCTTTTTCCCTTATTGCTTTTTATGTTCTTTCTTTGTTTTGTGCATTTGGTGTTCTGATT
ATTATGTGATGTGAGGTATTTCTCTTCTGGTCAAATCTATTTGGAGTTCTGTAGGCTTCT
TGTATGTTTATGGGCATCTCTTTCTTTAGGTTATGGATGTTTTCTTCTATAATTTTGTTG
AATATATCTACTGTCCCTTTAAGTTAGGAGCCTTCACTTTCTTCTATACCTGTTATCCTT
AGGTTTAATCTTCTCACTGGATTTCCTCGATGTTTTGGACTAGGAACTTTTTGCATTTTA
CATTATCTTTGACAGGTATTTCAATGTTTTCTATGGTATCTTCTGCCACTGAGATTCTCT
CTTCTAGCTCTTGTATAATGTTGGTGATGCTTGTACCTGTGACTCCTTGTTTCTTCCTTA
GGTTTTCTATCTCCAGGGTTGTCTCCCTTTGTGCTTTTTTTATTGCTTCTATTTCCATTC
TAAATCCTGGATGGTTTTGTTCAATTCCTTCACCTCTTTGGTTGTATTTTCCTGTAATTC

TTTCAGGGATTT[redacted]GTGTTTCCTCTTTAAGGGCTTCTACTTGT[redacted]ACTTGTGTTGTCCTG
TATTTCTTTAAGGTAGTTATTTATGTCCTTCTTGAAGTCCTCCATCATTATCAAAAAATG
TGATTTTTAAATATAAACCTTGCTTTTCTGGTGTGTTTGGATGTCAAGTATTTTCTTTGC
TGGGAGAACTGGGCTCTGATAATGCCAAGTTGTTTGATTTCTGTTGCTTAGTTTCCTGTT
CTTGCCTCTCGCCATTGGGTTTTCTCTGGTGTTTGCTTATCTTGCTGTTTCTGAGAGTGG
CTTGACACTCTTGTAGGCATCTGTGTCAGGCCTCCTGTAGAACTGTTTCCCTGTTTTCTT
TCAGCCTTTTCTGAGAACAGGTGCTCTGATCTCAGGTGTGTAGGCATTCCTGGTGACTAT
CTTTCAGCTTTAGGAGCAGGCAGGAATCAGAAGGGTCCTGTCCCTGACTGCTCCTAGATC
CTTGCACCCAGGGGGCACAGTTAGCACTAGGCAATTCCCTCTTGTGTAGGGAATGTGGGT
AGAGGATAGTCGCCTCTGATTTCTCAGGAATGTCTGCACTTCTGAAAGTCCAGCCCTCTC
CCCCACAGGATTTAGGTGCAGGGAGCTGTTTGACCACTTCAATTCAGTCCTGGGTGTAGA
CCAGAACCACAGGTAAAAAAGAATGACTTCATTAAATTAGCAGACAAATGGGTGGAACTA
GAAAATGTCATCCTGGGCTGGAGAGATGGCTCAGTGGTTCAGACCACTGGCTGCTCTTCC
AGAGGTCCTGAGTTCAATTCCCAACAACTATATGGTGGCTACCAACCATTACAATGAGAT
CAGATGCCCTCCTCTTGTGTATCTGAAGAGAGTGACAGTGTACTTACATACATAAAATAA
ATAAATAAATCTAAAAAAATGTTAAAAAA

SEQ ID NO:85

Rat T2R05 amino acid sequence

MLGAMEGVLLSVATSEALLGIVGNTFIALVNCMDCTRNKNLYNIGFILTGLAISRICLVW
ILITEAYIKIFSPQLLSPINIIELISYLWIITSQLNVWFATSLSIFYFLKIANFSHHIFL
WLKRRINIVFAFLIGCLLMSWLFSFPVVVKMVKDKKMLYINSSWQIHMKKSELIINYVFT
NGGVFLLFIIMLIVCFLLIISLWRHSKWMQSNESGFRDLNTEVHVKTIKVLLSFIILFIL
HLIGITINVICLLVPENNLLFVFGLTIAFLYPCCHSLILILANSRLKRCFVRILQQLMCS
EEGKEFRNT

SEQ ID NO:86

Rat T2R05 nucleotide sequence

AAGAGATTTCAGATACTACCACAAACATTTTTTAAATATATGTAAGTCTTTAAAGAAAGA
AGGGAAAGCCACTCCTTTATTGAGCAGCCAATAGATTGCCATCTTAAAATTCTGTGGCAG

AAGCTATTTTAA[redacted]TCTGCGAAGATGCTGGGTGCAATGGAAGG[redacted]TCCTCCTTTCAGTT
GCAACTAGTGAGGCTTTGCTTGGCATTGTAGGGAACACATTCATTGCACTTGTGAACTGC
ATGGACTGTACCAGGAACAAGAATCTCTATAATATTGGCTTCATTCTCACTGGCTTGGCA
ATTTCCAGAATCTGCCTCGTGTGGATCTTAATCACAGAGGCATACATAAAAATATTCTCT
CCACAGTTGCTGTCTCCTATCAACATAATTGAACTCATCAGTTATCTATGGATAATTACC
AGTCAATTGAATGTTTGGTTTGCTACCAGCCTCAGTATCTTTTATTTCCTCAAGATAGCA
AATTTTTCCCACCACATATTTCTCTGGTTAAAAAGAAGAATTAATATAGTTTTTGCCTTC
CTGATAGGGTGCTTACTTATGTCATGGCTATTTTCTTTCCCAGTAGTTGTGAAGATGGTT
AAAGATAAAAAAATGCTGTATATAAACTCATCTTGGCAAATCCACATGAAGAAAAGTGAG
TTAATCATTAACTATGTTTTCACCAATGGGGGAGTATTTTTACTTTTTATAATAATGTTA
ATTGTATGTTTTCTCTTAATTATTTCCCTTTGGAGACACAGCAAGTGGATGCAATCAAAT
GAATCAGGATTCAGAGATCTCAACACAGAAGTTCATGTGAAAACAATAAAAGTTTTATTA
TCTTTTATTATCCTTTTTATATTGCATTTAATTGGTATTACCATCAATGTCATTTGTCTG
TTAGTCCCAGAAAATAACTTGTTATTCGTGTTTGGTTTGACGATTGCATTCCTCTATCCC
TGCTGCCACTCACTTATCCTAATTCTAGCAAACAGCCGGCTGAAACGATGCTTTGTAAGG
ATACTGCAACAATTAATGTGCTCTGAGGAAGGAAAAGAATTCAGAAACACATGACAGTCT
GGAAGACAAACAATCAGAAATAGTAAGTGAAAAAAAAAAAAAAAAA

SEQ ID NO:87

Rat T2R06 amino acid sequence

EALVGILGNAFIALVNFMGWMKNRKITAIDLILSSLAMSRICLQCIILLDCIILVQYPDT
YNRGKEMRIIDFFWTLTNHLSVWFATCLSIFYFFKIANFFHPLFLWIKWRIDKLILRTLL
ACLILSLCFSLPVTENLADDFRRCVKTKERINSTLRCKLNKAGYASVKVNLNLVMLFPFS
VSLVSFLLLILSLWRHTRQMQLNVTGYNDPSTTAHVKATKAVISFLVLFIVYCLAFLIAT
SSYFMPESELAVIWGELIALIYPSSHSFILILGNSKLKQASVRVLCRVKTMLKGRKY

SEQ ID NO:88

Rat T2R06 nucleotide sequence

GTGAGGCCTTAGTAGGAATCTTAGGAAATGCATTCATTGCATTGGTAAACTTCATGGGCT
GGATGAAGAATAGGAAGATCACTGCTATTGATTTAATCCTCTCAAGTCTGGCTATGTCCA

GGATTTGTCTAC[redacted]TGTATAATTCTATTAGATTGTATTATATT[redacted]TGCAGTATCCAGACA
CTTACAACAGGGGTAAAGAAATGAGGATCATTGATTTCTTCTGGACGCTTACCAACCATT
TAAGTGTCTGGTTTGCCACCTGCCTCAGCATTTTCTATTTCTTCAAGATAGCAAACTTCT
TCCATCCTCTTTTCCTCTGGATAAAGTGGAGAATTGACAAGCTAATTCTGAGGACTCTAC
TGGCATGCTTGATTCTCTCCCTATGCTTTAGCCTCCCAGTCACTGAGAATTTGGCTGATG
ATTTCAGACGCTGTGTCAAGACAAAAGAAAGAATAAACTCTACTCTGAGGTGCAAATTAA
ATAAAGCTGGATATGCTTCTGTCAAGGTAAATCTCAACTTGGTCATGCTGTTCCCCTTTT
CTGTGTCCCTTGTCTCATTCCTTCTCTTGATTCTCTCCCTATGGAGACACACCAGGCAGA
TGCAACTCAATGTAACAGGGTACAATGATCCCAGCACAACAGCTCATGTGAAAGCCACAA
AAGCAGTAATTTCCTTCCTAGTTCTGTTTATTGTCTACTGCCTGGCCTTTCTTATAGCCA
CTTCCAGCTACTTTATGCCAGAGAGTGAATTAGCTGTAATTTGGGGTGAGCTGATAGCTC
TAATATATCCCTCAAGCCATTCATTTATCCTGATCCTTGGGAACAGTAAACTAAAACAGG
CATCTGTAAGGGTGCTTTGTAGAGTAAAGACTATGTTAAAGGGAAGAAAATATTAGCATC
ATGGATATATTTGAAGAAAAACTATCACTGTCTAAAGAAAAAGGATGACAAATCATTATC
TTTCATTCTTATATGAATATTGCTTTCATGCGGTAACATCTTTTAACAAACTTAAATCAA
ATGTTGGGAAATCTCATATACAGCAACTTTGCATGTCTCTCTGTCTATTTCCCTCTCCCT
TTGTACATAGTTGACATAAAAAAAGAATTTTCATGACAAAATTGTAATAAATAGCTACAG
AGGCAGCACATTTTCATAGTAAGTTCTGAATCACTCTTCCAAATGCAAAGCTGCCTGACA
AATTCAAAACAACTGTAACAGTATTTCACTGCTGTTTGCATTCTTTGGAAAAGCAGGTGG
TTTGTTCCTATGACCTGACTTGGAGTTTTCTTCTTACATCACTG

SEQ ID NO:89

Rat T2R07 amino acid sequence

MGSSLYDILTIVMIAEFIFGNVTNGFIVLTNCIAWLSKRTLSFIGWIQLFLAISRVVLIW
EMLLAWLKYMKYSFSYLAGTELRVMMLTWVVSNHFSLWLATILSIFYLLKIASFSRPVFL
YLKWRVKKVLLLILLGNLIFLMFNILQINTHIEDWMDQYKRNITWDSRVNEFVGFSNLVL
LEMIMFSVTPFTVALVSFILLIFSLWKHLQKMHLSSRGERDPSTKAHVNALRIMVSFLLL
YATYFISFFISLIPMAHKKGLDLMFSLTVGLFYPSSHSFILILGHSNLRHSSCLVITYLR
CKEKD

SEQ ID NO:90

Rat T2R07 nucleo[redacted] sequence

CAGTAGCAAAATTTTACTATGTTCATTGATATTATGTCAnGnCACTACGTAAGAAGGAAG
ACTTGAAAGAAAGCTTATCTGAGTTTTTAAGAATACATGGACATTTCAGCTTGGCAAATG
ACGAGCTGTGAATTTTTGTCATCTGGAC**ATGGGAAGCAGCCTGTATGATATCTTAACTAT**
TGTCATGATTGCAGAGTTTATATTCGGAAATGTGACCAATGGATTCATAGTGCTGACAAA
CTGTATTGCTTGGCTCAGTAAAAGAACTCTTTCTTTCATTGGTTGGATCCAGCTTTTCTT
GGCCATTTCCAGAGTGGTTTTGATATGGGAAATGTTACTAGCATGGCTGAAATATATGAA
GTATTCATTTTCATATTTGGCTGGCACAGAATTAAGGGTTATGATGTTGACCTGGGTAGT
TTCCAATCACTTTAGTCTCTGGCTTGCCACCATTCTAAGCATCTTTTATTTGCTCAAAAT
AGCTAGTTTCTCCAGACCTGTTTTCCTGTATCTGAAGTGGAGAGTAAAAAAAGTGCTCCT
GCTGATTCTTCTCGGAAATTTAATCTTCCTGATGTTCAATATATTACAAATCAACACTCA
CATAGAAGACTGGATGGATCAATATAAGAGAAATATAACGTGGGATTCCAGAGTGAATGA
ATTTGTGGGGTTTTCAAATCTGGTTTTATTGGAGATGATTATGTTCTCTGTAACACCATT
CACCGTGGCTCTGGTCTCCTTCATCCTGTTAATCTTCTCTTTATGGAAACATCTCCAGAA
GATGCATCTCAGTTCCAGAGGGGAACGAGACCCTAGCACAAAAGCCCATGTGAATGCCCT
GAGAATTATGGTCTCCTTCCTCTTACTCTATGCCACTTACTTCATATCCTTTTTTATATC
ATTAATTCCTATGGCACATAAAAAAGGACTAGATCTTATGTTTAGCCTAACTGTTGGACT
TTTCTACCCTTCAAGCCACTCATTTATCTTGATTTTGGGACATTCTAATCTAAGGCATTC
CAGTTGTCTGGTGATAACCTATCTGAGATGTAAGGAAAAGGATTAGAAATTCACTATTCC
ATAAGGCAGTTAAACCACATGCTATTAGGTATACTCAGTGCTAGATCCCTAGGCAAGCAT
TAACATTAAAAATATATAATTTCTAGATTCTTCTATTTGTGATAAACCACTCACTTAGAA
TAATGCTAAAGTAGCGTGATGTTGTATATAAGTGTAAGAATAAAATGTAATTAATTTAGT
TTAGGCACAATAACATATGTCTACTAAGTAAAAACTAGGCAGGCTGCTACACGCATATTA
GAATCCAGGCTGAGGTATATAGACTCAAGAAATACTGTGGAATAAAGATTTTAATTTTCA
TTCTATTGTGAGTTATGTGAAATCAATGCCATTAAAGGCATACACAAGATTTTCACACAC
TGAAACAACTTCTTGCATTTTGTCATATTGTATTGGAAGTAAATTGGAGATAAACTTAAT
ATCAATAAATTACAAAATGTAAACATAAACAGGGTGATTAAAAATTAGCCTCTAGGTCCT
GGGGAAATGATTCaAGTAAAGTGCTTTCTTTTCAAATAGGAGAATCTGATTGTAAATCAT
CTAAAAGTCTGGCATAAAATGTCAATGAAAATTGTATGTAAAATATAGCTATgGCmAAGA
GCACCmAAGAAAAGAAAATTTTTGCCTTTGAAACCCAGTAATTGATATCCTTTAAAAAAG
CAGTTACATATTTTTCTGTTTAAGATTTTGTCAAAGGGTAGCTTTGACAACTAATATAAG
CTGAGGAAGGTAGCAAGTGTGAAGTCAGCTAATGGGGTCAGTCAAGTGCTGTTAGCAGCA
GATGGAGGCCACTGCTGAATTTAGCAGGCAATTTACAGGGTGAGCACTGCTAGTGCTGAC

AGAAGAAAAACT[redacted]GAAATTTTAACTCTTTAGGGTCTGGTGAG[redacted]GAAAAAGAGAGAAA
ATCGCATATATATATATATATATATATATATATATATATATATATATATATATATATATA
TCATGGAAGCTCTAACAAGTTGACTCAAACAACTTTATGATGTTTTTAGGCCCTTTTATT
TTAATGTCAGTGAATTAGGTGTGGTACAGCAATATTGCTACTTTTAAATTCAAAGCAGTT
GTTTTATATATTATTCATTATATAAGCTAATTATAAGTTTAAATCAAAAGGTTTATTTGT
CCATGATTTTACTTTATCATTGGGCACACCTGTGCTCTCATCCTTGGGCTTGACCTAGAA
TGAAAGTTTATCCTTGATCATATGTCTGTCACAAGACTACTTCTCTTCCTATAGTAGTTT
ATGTACTTACAATATACAAAAGTTTATTGAATTCCTTTTATCACTTATGCAGCCTTTTCT
TACTATTCTATTCTATTCTATTCTATTCTATTCTATTCTATTCTATTCTATTCTATTCTA
TTCTATTCTATTCTATTCTAGAATCTAACCTATACATTCATTTCTGGCAAAACAACTTAT
ATCATCTCCTTAATTATTTTATCAATTAATCTAACATCCTGAAGTTATTTAAATCTAATA
TAAGGACTCTGTAAAGTCACAAATTTATTTATACTTCACAAAATTCATTATTTTATGGAA
CTGCAGCATTGCCTGGGCCAGGAGTCACAAGAGTTCCAGAGTTGACTTTATTGGCATCTG
CCTGGCTAACTGAAGGATCAGTTTTCTGTGTACAATAATTTTGTGTATCTCTTTTGATGC
AAGATATGAAAAATAATTTCAGTCTAAAAGTGTCCTTAAATTTGAAACTCTCTGGCCAGA
ATCTAACTATTGATGACCAGTTTGCACCATGGACTCAGTGTCTTCTATTGCTTTAAAATA
AGCAACATCTTGAATGCTTTTCTTGTGTATTAGGCAAATAATTAACAACATGTTTCTATG
ATTGTCTCAATAACAATACTATATTTCTCACAGTTTTTAATTTTTATGGCAAAGTTGGCT
AATAAGAATTTTTTTCAAATTATCAAACGTGAAGAAAACTTGACATTTTATTTCATGGAG
ATTCTAAATGTTTTCTTAGCATATTGCCTTTTTACTAACTTGATTTTTATCATGTTTTGG
TAGTATTTCTAATTTTCCTTTTTTTCTAAGTATGTTATGTAGTAACACCAGGAGAATGAA
ACAAATGACATTTATACTAAGGATGTGACAAATAAGGCCCAAAGAAAGTTTTGAAAATCA
TGATCTCATTTCTATTCTTCTTTATTAAGTATAGCATAAGCAAAATTCTGATGGTGGTCT
TGGCCCATATCTTTGAACACAGTGTAGTGGTGAAGACTTTTTCAAATATTATGTCATATT
TGTACCCATCTCTGTACCTATTTCTTCTGATTTCATGAGGAAAAAATGAGGAAGGGTTTG
TTTGTGTGCTGGAGCAGCTGAAGTGGACCAAGGGGCAGGAATTCTCTCTGTTCGGTCCTA
GTGTGACTGATGATGCTCTCATTGAAAAACAGGAAGAAGAAGAAAGACTTTATATGCACC
ATTCACTCCTTCCCCCTCCTACATTCCACCTCCCTCTTGAAAGAGTGTCTATCTATATAG
ATATAGCTATCCTGAAATCCATTAAGTAGACCTGACTGGCTTAAATCTCACAGAAATTCA
CCTACCTTTTCCATGATTGCTGAAATTAAAGACATGTGCCGACATATTGGGCACATTCAG
ACCTTTTGCCAACTGTCTTTCAACTCATTTGGACCTACTGAGAAGTATTCAAAATATTTG
GTTGTTTTAAATAAAAGGAAAGTGGGTCTATATTACTTGAATTGGATAGAGAAATTTTCA
CTTACAAGTGATATTGAAAATGGGGGAGAATGTATTTTAGCATAAGCACCAGAACACAAA

GCAATTCTTGTT[redacted]AACTTTATCGATAAATTGGATAAATGTTA[redacted]AAGAAAAAATAAAA
TATACGAACTATTATGAAAAAAAAAAAAAAAAA

SEQ ID NO:91

Rat T2R08 amino acid sequence

MEPVIHVFATLLIHVEFIFGNLSNGLIVLSNFWDWVVKRKLSTIDKILLTLAISRITLIW
EMYACFKIVYGSSSFIFGMKLQILYFAWILSSHFSLWFATALSIFYLLRIANCSWKIFLY
LKWRLKQVIVGMLLASLVFLPGILMQRTLEERPYQYGGNTSEDSMETDFAKFTELILFNM
TIFSVIPFSLALISFLLLIFSLWKHLQKMQLSSRGHGDPSTKAHRNALRIMVSFLLLYTS
YFLSLLISWIAQKHHSKLVDIIGIITELMYPSVHSFILILGNSKLKQTSLWILSHLKCRL
KGENILTPSGKPIN

SEQ ID NO:92

Rat T2R08 nucleotide sequence

CTGCAGGTTGGTGATCCAGTAATGAGCAGCACTGTTATATCTCAGGCTTTCTAAGATCAT
GGAACCTGTCATTCACGTCTTTGCCACTCTACTAATACATGTGGAGTTCATTTTTGGGAA
TCTGAGCAATGGATTAATAGTGTTGTCAAACTTCTGGGACTGGGTCGTTAAACGAAAACT
TTCCACAATTGATAAAATTCTTCTTACATTGGCAATTTCAAGAATCACTCTCATCTGGGA
AATGTATGCTTGTTTTAAAATTGTATATGGTTCATCTTCATTTATATTTGGGATGAAGTT
ACAAATTCTTTATTTTGCCTGGATCCTTTCTAGTCACTTCAGCCTCTGGTTTGCCACAGC
TCTCAGCATCTTTTACTTACTCAGAATAGCTAACTGCTCCTGGAAGATCTTCCTGTATCT
GAAATGGAGACTTAAACAAGTGATTGTGGGGATGTTGCTGGCAAGCTTGGTGTTCTTGCC
TGGAATCCTGATGCAAAGGACTCTTGAAGAGAGGCCCTATCAATATGGAGGAAACACAAG
TGAGGATTCCATGGAAACTGACTTTGCAAAGTTTACAGAGCTGATTCTTTTCAACATGAC
TATATTCTCTGTAATACCATTTTCATTGGCCTTGATTTCTTTTCTCCTGCTAATCTTCTC
TTTGTGGAAACATCTCCAGAAGATGCAGCTCAGTTCCAGAGGACATGGAGACCCTAGCAC
CAAGGCCCACAGAAATGCTTTGAGAATTATGGTCTCCTTCCTCTTGCTCTACACTTCATA
TTTCCTGTCTCTTCTTATATCATGGATTGCTCAGAAGCATCACAGTAAACTGGTTGACAT
TATTGGTATTATTACTGAACTCATGTATCCTTCAGTCCACTCATTTATCCTGATTCTAGG
AAATTCTAAATTAAAGCAGACTTCTCTTTGGATACTGAGTCATTTGAAATGTAGACTGAA

AGGAGAGAATAT[redacted]TAACTCCATCTGGCAAACCAATTAACTAG[redacted]TTATATATTCTGTA
TTGCAAACAAATCAGTGAGTTAGTGGTTCAAGGATTCCATCCTTGACTTATTGTATCATG
GAAGTCATATAGGGAGAGGCTGAACAAGCTATCTTCTGTAAATTGGCAAGGGTTGCATAT
AGTACTGGTACTGGGACACCATCCAACCATAAAACCTTCTAACCATAACCTACCTGACTG
CAAGATATGCTGGGACAATGGTGGCTCAGAGATTTTGGGACTGGCCAACCAATGTCTATT
CTTTCTTGAGGCTCACTCAATAAGGAGGCCATGCCCAACTCGTCcTGGATGGCCAGGAAC
CAGAATCTCTGATGGsCCAATGATCTATGGnAGAACCCAGCATTACTGGGAAAAAAGAAT
AATCACTTTGATGAATGGTCAAATATTTCCTAAATATATTCTGATACACTTGTACATCAT
TTCTCTTTCCCAATCATCATCACAGGGACTTCTCCCCAGCACCTGATGGGAACAGATACC
AAAATCTACAGCCAAATACTAAATGCAGGTTGGGGAACTCCACAAAAGACTGGAAGGAAG
TACTGTGAGAGCCAGAGTGGTCCAGAACACTAGGAGAACACAGAACATCGAATTAACTAA
GCAGCACTCATAGGGTTAATGTAAAATAAAGCAGCAGTCACATAGACTGCACAGGTGTAC
TCTAGATCCTCTGCATATATGTTGTGGTTGTCAAACTTGGGAGTTTTGTTGGACTAATAA
CAATGTGAATAAGTAAGTCTCTGACACTTATTCCCGCTCTTGGAACCCTTTTCCACATTT
TGTATTGTCTTACCACCTTGATATGAAGGTTTCTGAATAGTCCAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAA

SEQ ID NO:93

Rat T2R09 amino acid sequence

MLSAAEGILLSIATVEAGLGVLGNTFIALVNCMDWAKNKKLSKIGFLLFGLATSRIFIVW
ILILDAYAKLFFPGKYLSKSLTEIISCIWMTVNHMTVWFATSLSIFYFLKIANFSHYIFL
WLKRRTDKVFAFLLWCLLISWAISFSFTVKVMKSNPKNHGNRTSGTHWEKREFTSNYVLI
NIGVISLLIMTLTACFLLIISLWKHSRQMQSNVSGFRDLNTEAHVKAIKFLISFIILFIL
YFIGVAVEIICMFIPENKLLFIFGLTTASVYPCCHSVILILTNSQLKQAFVKVLEGLKFS
ENGKDLRAT

SEQ ID NO:94

Rat T2R09 nucleotide sequence

GGACACTGCAGCAGATCTGCTATAGAATAACAGATACAAACATAGCAACCTGCAGAG**ATG**
CTCAGTGCAGCAGAAGGCATCCTTCTTTCCATTGCAACTGTTGAAGCTGGGCTGGGAGTT

TTAGGGAACACAT●TATCGCCCTGGTTAACTGCATGGATTGGGC●AGAACAAGAAGCTC
TCTAAGATTGGTTTCCTTCTCTTTGGCTTAGCAACTTCCAGAATTTTTATTGTATGGATA
TTAATTTTAGACGCATATGCAAAGCTATTCTTTCCGGGGAAGTATTTGTCTAAGAGTCTG
ACTGAAATCATCTCTTGTATATGGATGACTGTGAATCACATGACTGTCTGGTTTGCCACC
AGCCTCAGCATCTTCTATTTCCTAAAAATAGCAAATTTTTCCCACTATATATTTCTCTGG
TTAAAGAGGAGAACTGATAAAGTATTTGCCTTTCTCTTGTGGTGTTTATTAATTTCATGG
GCAATCTCCTTCTCATTCACTGTGAAAGTGATGAAGAGCAATCCAAAGAATCATGGAAAC
AGGACCAGTGGGACACATTGGGAGAAGAGAGAATTCACAAGTAACTATGTTTTAATCAAT
ATTGGAGTCATTTCTCTCTTGATCATGACCTTAACTGCATGTTTCTTGTTAATTATTTCA
CTTTGGAAACACAGCAGGCAGATGCAGTCTAATGTTTCAGGATTCAGAGATCTCAACACT
GAAGCTCATGTGAAAGCCATAAAATTTTTAATTTCATTTATCATCCTTTTCATCTTGTAC
TTTATAGGTGTTGCAGTAGAAATCATCTGCATGTTTATCCCAGAAAACAAACTGCTATTT
ATTTTTGGTTTGACAACTGCATCCGTCTATCCCTGCTGTCACTCAGTCATTCTAATTCTA
ACAAACAGCCAGCTGAAGCAAGCCTTTGTAAAGGTACTGGAGGGATTAAAGTTCTCTGAG
AACGGAAAAGATCTCAGGGCCACATGAGTCTGGAACAGAAATGGGTAGTCTGGAATAATT
GTAAGGAAGTCGTAGAAGGTCTTTTTCATTTGTACAGTGCTCTTACCTTGTTTTTGAGGA
GATGTAAACTTTTTTATTTTTATTTTTTATCCTATGTGAATAAGTGTGTGTGTGTGTGTG
TGTGTTTATGTGTGTGTGTATATATGTCTATGTGTGTTTTAGGAGGTTTAAGAGGGAAGA
GGGAATAGAGGTATGTTGGTGTTTTTAACATGGATATTCACAGGCCAAGGAACTTGTTCT
CTCCTTTTACCTTAGGGTAGTGTCCTTTGTGGCTGTCACTCTGACAGTCTACACTAGTTG
AACTAAGAGCTTTTAGCCAGTTCACTTGTCTAAACCTCCCTTCTCATGGTAGCAGTGTTC
TGATTACAGAATCATGCTGTCACATACAGCTTTTTAACAAGGTTCCCATAGACAGAATTC
ATGTCAAACGGAATGCACAGCTGTCACTCTTACCCACCGATCTCTCTTGCCAGCCCATTC
CTATTGACTTTAAACTGTAGTATTAAACTTTACTGAAATCTTCTGCAACCAGTCTGACTA
TGTCTCTTGAAATCACATGATATGGTGGAATTTTAATGCCATGTGAAAATTTGTTTGTTC
AGTTAGTTTCCTACTCTGCCAAATCATTCTCTTACACTTGGCAGAAAAAAACCATCAACT
GTAGACTATTTTGTGTAAAGACTAATACAGATAGAATAAGTATCTTAATCAAGATGTCAT
TGTGATTATCCTAATTTCCCCAGAGCACTGGTTCCCTTTCCCCAGAAAGACTCACAAAGG
AACTGAGGCAAACAGTTGTGGTCACTCTTGATATTTACCAGTTGAAACTGAAGAACAGTG
TTTCCTTTCTGTTCAGTTTTACTACTTACAGTTACTTTATTTCATCCATTAAATCCCAAA
GTGCTTATTAATAGTAGATATTTGATGAAGCAACAATGGTTATAAGAGTGGATGTGGATC
TATGACAAAGATCTAGAGAAACAGACTATTTGTGAAAGATGGATGAAAGCCCTGATGAAA
GGATTCTTCATGGTCTTTGACCCCAGGGAGTTTTGAAATCAAGCAGCCACAGATCAAAGA
GAGCTGAGAAGAGGTTCTCCTGAAGAAAATATCCAAACACATGGTGCCAGCCAAAGCAGA

AAATAGTGGAC[redacted]TCAGTCCAGGACCTGAATGAGGTAGACAA[redacted]TCCTGTTAAGGGTTG
GAACAAATATATAGATATGGTCATTCATATACAGAAACCTACAGGCGTGTTTGAACTCTT
GGTTTCTCAGTAATCAATTCTTAAATCTTTTTTAGAATGGATTTTTTATCATCATTCATG
ATCTCTCAGCAGAGTCTGCAGGGGCTAAGAGACACACTAAGAGTATCTGGAGGGGGGAGT
GTCTTCCTGCTCTATCAACCCCTAAAGTCATATATAACAATACAAAATTCCACATTAGTT
AAGTTCTTTTTTTTACATCTTTATTAAATTGGGTATTTCTTATTTACATTTCAAATGTGA
TTCCCTTTCCTGGTTTCCAGGCCAATATCCCCCTAACCTCTCCCCTTCTATGTGGGTATT
CCCTCGTGCCGAATTC

SEQ ID NO:95

Rat T2R10 amino acid sequence

MFLHTIKQRDIFTLIIIFFVEITMGILGNGFIALVNIVDWIKRRRISSVDKILTTLALTR
LIYAWSMLIFILLFILGPHLIMRSEILTSMGVIWVVNNHFSIWLATCLGVFYFLKIANFS
NSLFLYLKWRVKKVVLM

SEQ ID NO:96

Rat T2R10 nucleotide sequence

CCCGGGCTGCAGGATTCGGCACGAGAATGAAAACTTTTGCTCTACTATTTTGCTGTTCTG
TGATACCACAGACCATAAAACAATCGAGCCAAGGGATCAAGAGCTGAAACTTCAGAAAGT
GGGAATCAAATTTCCTTCCTGATAGGTTAGCTTATGAGAATTCAGCATCTTATTCAACTT
CAGAAAATTGGATATAAGATACAGTGTCTGGATGAAGCCGAATTGATCTATTTGGGGAGA
AAAAACGCCAACATTTATAATAAGGTTTTATGAGACAGTTCCTGGGAAATTTGGATATTT
CCTAGTTAGTAATGTGTAAATGGGATTTTAAAACATGATTATTTTGTATTTTTAACAACC
AACATGAGGAGCTTTTTAAATGCCACTTAGACATTATAAACTGAAGC**ATGTTCTTACACA**
CAATAAAGCAACGTGATATTTTTACTTTGATAATCATATTTTTTGTGGAAATAACAATGG
GAATCTTAGGAAATGGATTCATAGCACTAGTGAACATTGTGGACTGGATCAAGAGAAGAA
GGATTTCTTCAGTGGATAAGATTCTCACTACCTTGGCCCTTACCAGACTCATTTATGCGT
GGTCTATGCTCATTTTTATATTGTTATTCATACTGGGCCCGCATTTGATTATGAGATCAG
AAATACTTACATCAATGGGTGTTATCTGGGTGGTGAACAATCACTTCAGCATCTGGCTTG

CTACATGCCTCGG[redacted]TCTTTTATTTTCTCAAGATAGCCAATTTT[redacted]TAACTCTTTGTTTC
TTTACCTAAAGTGGAGAGTTAAAAAAGTGGTTTTAATG

SEQ ID NO:97

Rat T2R11 amino acid sequence

GSGNGFIVSVNGSHWFKSKKISLSDFIITSLALFRIFLLWIIFTDSLIIVFSYHAHDSGI
RMQLIDVFWTFTTHFSIWLISCLSVFYCLKIATFSHPSFL*LKSR

SEQ ID NO:98

Rat T2R11 nucleotide sequence

GGATCCGGAAACGGTTTTATCGTGTCAGTCAATGGCAGCCATTGGTTCAAGAGCAAGAAG
ATTTCTTTGTCTGACTTCATCATTACCAGCTTGGCCCTCTTCAGGATCTTTCTGCTGTGG
ATCATCTTTACTGATAGCCTCATAATAGTGTTCTCTTACCACGCCCACGACTCAGGGATA
AGGATGCAACTTATTGATGTTTTCTGGACATTTACAACCCACTTCAGTATTTGGCTTATC
TCCTGTCTCAGTGTTTTCTACTGCCTGAAAATAGCCACTTTCTCCCACCCCTCATTCCTG
TAGCTCAAATCTAGA

SEQ ID NO:99

Rat T2R12 amino acid sequence

MLSTVSVFFMSIFVLLCFLGILANGFIVLMLSREWLWRGRLLPSDMILLSLGTSRFCQQC
VGLVNSFYYSLHLVEYSRSLARQLISLHMDFLNSATFWFGTWLSVLFCIKIANFSHPAFL
WLKWRFPALVPWLLLGSILVSFIVTLMFFWGNHTVYQAFLRRKFSGNTTFKEWNRRLEID
YFMPLKLVTTSIPCSLFLVSILLLINSLRRHSQRMQHNAHSLQDPNTQAHSRALKSLISF
LVLYALSYVSMVIDATVVISSDNVWYWPWQIILYLCMSVHPFILITNNLKFRGTFRQLLL
LARGFWVT

SEQ ID NO:100

Rat T2R12 nucleotide sequence

GTGTGAGGGACTGTGGGTAGGGGCTGGGAGGAGGCCAGGAACCAAGGCAACCAGTGGTGA
CAGGAGGGGCTGAAATGCTATCAACTGTATCAGTTTTCTTCATGTCGATCTTTGTTCTGC
TCTGTTTCCTGGGAATCCTGGCAAACGGCTTCATTGTGCTGATGCTGAGCAGGGAATGGC
TATGGCGCGGTAGGCTGCTCCCCTCAGACATGATCCTCCTCAGTTTGGGCACCTCCCGAT
TCTGCCAGCAGTGCGTTGGGCTGGTGAACAGTTTCTACTATTCCCTCCACCTTGTTGAGT
ACTCCAGGAGCCTTGCCCGTCAACTCATTAGTCTTCACATGGACTTCTTGAACTCAGCCA
CTTTCTGGTTTGGCACCTGGCTCAGCGTCCTGTTCTGTATCAAGATTGCTAACTTCTCCC
ATCCTGCCTTCCTGTGGTTGAAGTGGAGATTCCCAGCATTGGTGCCTTGGCTCCTACTGG
GCTCTATCTTGGTGTCCTTCATCGTAACTCTGATGTTCTTTTGGGGAAACCACACTGTCT
ATCAGGCATTCTTAAGGAGAAAGTTTTCTGGGAACACAACCTTTAAGGAGTGGAACAGAA
GGCTGGAAATAGACTATTTCATGCCTCTGAAACTTGTCACCACGTCAATTCCTTGCTCTC
TTTTTCTAGTCTCAATTTTGCTGTTGATCAATTCTCTCAGAAGGCATTCACAAAGAATGC
AGCACAATGCTCACAGCTTGCAAGACCCCAACACCCAGGCTCACAGCAGAGCCCTGAAGT
CACTCATCTCATTTCTGGTTCTTTACGCGCTGTCCTATGTGTCCATGGTCATTGACGCTA
CAGTTGTCATCTCCTCAGATAACGTGTGGTATTGGCCCTGGCAAATTATACTTTACTTGT
GCATGTCCGTACATCCATTTATCCTTATCACTAATAATCTCAAGTTCCGAGGCACCTTCA
GGCAGCTACTCCTGTTGGCCAGGGGATTCTGGGTGACCTAGAAGGTTTGGTCTCTTTATC
TGTACCCTTTGAAGAGACTTAGGTGAGGGTGACTTCCCTTGGAAGTGATCTCATCTACAT
GGAAATGTCTTTGTAGGCTGACATGGGGTCATACTATGTGGTTCCTCCTTGGGAAAGAGG
AGAAGAAAATACAGGGATTCTGAGCGTTCTTCCTTATCTTGGGATATTATGAAAATGGAC
ATTCTGAATCCTGAACCAGTATTGATCTGAAGTGCAAAGTACAATATGCCTGTTCCCTTC
ATGTCTGCTATCCTCTTGGTACTTATTAATTCCCT

SEQ ID NO:101

Rat T2R13 amino acid sequence

MCGFPLSIQLLTGLVQMYVILIIAVFTPGMLGNVFIGLVNYSDWVKNKKITFINFILICL
AASRISSVLVVFIDAIILELTPHVYHSYSRVKCSDIFWVITDQLSTWLATCLSIFYLLKI
AHFSHPLFLWLKWRLRGVLVGFLLFSLFSLIVYFLLLELLSIWGDIYVIPKSNLTLYSET
IKTLAFQKIIVFDMLYLVPFLVSLASLLLLFLSLVKHSQNLDRISTTSEDSRAKIHKKAM

KMLLSFLVLFIIH[redacted]CMQLSRWLFFLFPNNRSTNFLLLTLNIFP[redacted]HTFIIILGNSKLRQ
RAMRVLQHLKSQLQELILSLHRLSRVFTMEIA

SEQ ID NO:102

Rat T2R13 nucleotide sequence

GGGATTCAGTTGGATAAGAGAAAAGTCAAAACCCTAAGACTAAGAATTTCCTTAAGTAGA
TATCAATTTCTATCCATTGGAAGGAGTTTCCAATCACACTGAAATTACAATAAAAAAGGA
GCAAGATAACTATGGGAAAGGATGATTTTCGGTGGATGTTTGAGAACTGAGCAGCAAGGC
AAATTGATAG**ATGTGTGGATTCCCTCTTTCTATTCAACTGCTTACTGGATTGGTTCAAAT**
GTACGTGATATTGATAATAGCAGTGTTTACACCTGGAATGCTGGGGAATGTGTTCATTGG
ACTGGTAAACTACTCTGACTGGGTAAAAAACAAGAAAATCACCTTCATCAACTTCATCCT
GATCTGTTTGGCAGCGTCCAGAATCAGCTCTGTGTTGGTGGTATTTATTGATGCAATCAT
CCTAGAACTAACTCCTCATGTCTATCATTCTTACAGTCGAGTGAAATGCTCTGATATATT
CTGGGTTATAACTGACCAGCTGTCAACGTGGCTTGCCACCTGCCTCAGCATTTTCTACTT
ACTCAAAATAGCCCACTTCTCCCATCCCCTTTTCCTTTGGTTGAAGTGGAGATTGAGAGG
AGTGCTTGTTGGTTTTCTTCTATTTTCTTTGTTCTCATTGATTGTTTATTTTCTACTCCT
GGAATTACTGTCTATTTGGGGAGATATTTATGTGATCCCTAAAAGCAATCTGACTTTATA
TTCAGAAACAATTAAGACCCTTGCTTTTCAAAAGATAATTGTTTTTGATATGCTATATTT
AGTCCCATTTCTTGTGTCCCTAGCCTCATTGCTCCTTTTATTTTTATCCTTGGTGAAGCA
CTCCCAAAACCTTGACAGGATTTCTACCACCTCTGAAGATTCCAGAGCCAAGATCCACAA
GAAGGCCATGAAAATGCTATTATCTTTCCTCGTTCTCTTTATAATTCACATTTTTTGCAT
GCAGTTGTCACGGTGGTTATTCTTTTTGTTTCCAAACAACAGGTCAACTAATTTTCTTTT
GTTAACATTAAACATCTTCCCATTATCTCATACATTCATTATCATCCTGGGAAACAGCAA
GCTTCGACAAAGAGCAATGAGGGTCCTGCAACATCTTAAAAGCCAACTTCAAGAGTTGAT
CCTCTCCCTTCATAGATTGTCCAGAGTCTTCACTATGGAAATAGCTTAAAGGGGAGACTT
GGAAGGTCACTGGTAACTTGTTCTTCCGCTGAGTTCTGTTAAGTAATGCTGGACATATAT
GAACTATCCCTAGTGCATACTGATATT

SEQ ID NO:103

Rat T2R14 amino acid sequence

VANIMDWVKRRKL●VDQLLTVLAISRITLLWSLYILKSTFSMV●FEVAIPSTRLTNLV
WIISNHFN

SEQ ID NO:104

Rat T2R14 nucleotide sequence

CTGTGGCAAACATAATGGATTGGGTCAAGAGAAGGAAGCTCTCTGCAGTGGATCAGCTCC
TCACTGTGCTGGCCATCTCCAGAATCACTCTGTTGTGGTCATTGTACATACTGAAATCAA
CATTTTCAATGGTGCCAAACTTTGAGGTAGCTATACCGTCAACAAGACTAACTAATCTTG
TCTGGATAATTTCTAACCATTTTAAT

SEQ ID NO:105

Mouse T2R01 amino acid sequence

MQHLLKTIFVICHSTLAIILIFELIIGILGNGFMALVHCMDWVKRKKMSLVNKILTALAI
SRIFHLSLLLISLVIFFSYSDIPMTSRMTQVSNNVWIIVNHFSIWLSTCLSVLYFLKISN
FSNSFFLYLKWRVEKVVSVTLLVSLLLLILNILLINLEISICIKECQRNISCSFSSHYYA
KCHRQVIRLHIIFLSVPVVLSLSTFLLLIFSLWTLHQRMQQHVQGGRDARTTAHFKALQT
VIAFFLLYSIFILSVLIQNELLKKNLFVVFCEVVYIAFPTFHSYILIVGDMKLRQACLPL
CIIAAEIQTTLCRNFRSLKYFRLCCIF

SEQ ID NO:106

Mouse T2R01 nucleotide sequence

AGCTGTGCGTGAGCAAAGCATTTCTTGTCTGCCACTTCTGAGCTGTGTGAGGAGACACAT
TATCACGGAAAGAGATTCAGACTCTGTCGCTGTCAAACCTGTATGTTTGCTCCTCTTTTA
CTGTGAAGGCAGAGTTACGAAAAAAAATGTTATGAGAACCAACTCAGAAATTGACAAAAA
TTTTCTAAATGTCATTTTTAAAAATTATATTTCAAATGGAAATGTGAGCAAATCTTTATA
ACTAATATATAAA**ATGCAGCATCTTTTAAAGACAATATTTGTTATCTGCCATAGCACACT**
TGCAATCATTTTAATCTTTGAATTAATAATTGGAATTTTAGGAAATGGGTTCATGGCCCT
GGTGCACTGTATGGACTGGGTTAAGAGAAAGAAAATGTCCTTAGTTAATAAAATCCTCAC

TGCTTTGGCAATC[redacted]CAGAATTTTTCATCTCAGTTTATTGCTTA[redacted]AGTTTAGTCATATT
CTTTTCATATTCTGATATTCCTATGACTTCAAGGATGACACAAGTCAGTAATAATGTTTG
GATTATAGTCAATCATTTCAGTATCTGGCTTTCTACATGCCTCAGTGTCCTTTATTTTCT
CAAGATATCCAATTTTTCTAACTCTTTTTTTCTTTATCTAAAGTGGAGAGTTGAAAAAGT
AGTTTCAGTTACACTGTTGGTGTCATTGCTCCTCCTGATTTTAAATATTTTATTAATTAA
CTTGGAAATTAGCATATGCATAAAGGAATGTCAAAGAAACATATCATGCAGCTTCAGTTC
TCATTACTATGCAAAGTGTCACAGGCAGGTGATAAGGCTTCACATTATTTTCCTGTCTGT
CCCCGTTGTTTTGTCCCTGTCAACTTTTCTCCTGCTCATCTTCTCCCTGTGGACACTTCA
CCAGAGGATGCAGCAGCATGTTCAGGGAGGCAGAGATGCCAGAACCACGGCCCACTTCAA
AGCCCTACAAACTGTGATTGCATTTTTCCTACTATATTCCATTTTTATTCTGTCTGTCTT
AATACAAATATGAATTACTGAAGAAAAATCTTTTCGTTGTATTTTGTGAGGTTGTATATA
TAGCTTTTCCGACATTCCATTCATATATTCTGATTGTAGGAGACATGAAGCTGAGACAGG
CCTGCCTGCCTCTCTGTATTATCGCAGCTGAAATTCAGACTACACTATGTAGAAATTTTA
GATCACTAAAGTACTTTAGATTATGTTGTATATTCTAGACAAAAATTAACTGATACAAAT
GTCTTTTGTATTTTTCATTTTAAATATCCTTTAATTTTGACTGCATGAAATTGATTTCTG
CTTGCAATTATCACTGATTAAAACTATTAATAATTTAACTAGTTGTATACAAGG

SEQ ID NO:107

Mouse T2R02 amino acid sequence

MESVLHNFATVLIYVEFIFGNLSNGFIVLSNFLDWVIKQKLSLIDKILLTLAISRITLIW
EIYAWFKSLYDPSSFLIGIEFQIIYFSWVLSSHFSLWLATTLSVFYLLRIANCSWQIFLY
LKWRLKQLIVGMLLGSLVFLLGNLMQSMLEERFYQYGRNTSVNTMSNDLAMWTELIFFNM
AMFSVIPFTLALISFLLLIFSLWKHLQKMQLISRRHRDPSTKAHMNALRIMVSFLLLYTM
HFLSLLISWIAQKHQSELADIIGMITELMYPSVHSCILILGNSKLKQTSLCMLRHLRCRL
KGENITIAYSNQITSFCVFCVANKSMR

SEQ ID NO:108

Mouse T2R02 nucleotide sequence

CAGCACAGTGAAAAACTCATGGGCCACTTGGTCACCCAGGGACAGGCGACGCTGTTATAT
GCCAAGCTTTCTATGAAC**ATGGAATCTGTCCTTCACAACTTTGCCACTGTACTAATATAC**

GTGGAGTTTATT[redacted]TGGGAATTTGAGCAATGGATTCATAGTGTT[redacted]CAAACTTCTTGGAC
TGGGTCATTAAACAAAAGCTTTCCTTAATAGATAAAATTCTTCTTACATTGGCAATTTCA
AGAATCACTCTCATCTGGGAAATATATGCTTGGTTTAAAAGTTTATATGATCCATCTTCC
TTTTTAATTGGAATAGAATTTCAAATTATTTATTTTAGCTGGGTCCTTTCTAGTCACTTC
AGCCTCTGGCTTGCCACAACTCTCAGCGTCTTTTATTTACTCAGAATAGCTAACTGCTCC
TGGCAGATCTTTCTCTATTTGAAATGGAGACTTAAACAACTGATTGTGGGGATGTTGCTG
GGAAGCTTGGTGTTCTTGCTTGGAAATCTGATGCAAAGCATGCTTGAAGAGAGGTTCTAT
CAATATGGAAGGAACACAAGTGTGAATACCATGAGCAATGACCTTGCAATGTGGACCGAG
CTGATCTTTTTCAACATGGCTATGTTCTCTGTAATACCATTTACATTGGCCTTGATTTCT
TTTCTCCTGCTAATCTTCTCTTTGTGGAAACATCTCCAGAAGATGCAGCTCATTTCCAGA
AGACACAGAGACCCTAGCACCAAGGCCCACATGAATGCCTTGAGAATTATGGTGTCCTTC
CTCTTGCTCTATACCATGCATTTCCTGTCTCTTCTTATATCATGGATTGCTCAAAAGCAT
CAGAGTGAACTGGCTGATATTATTGGTATGATAACTGAACTCATGTATCCTTCAGTCCAT
TCATGTATCCTGATTCTAGGAAATTCTAAATTAAAGCAGACTTCTCTTTGTATGCTGAGG
CATTTGAGATGTAGGCTGAAAGGAGAGAATATCACAATTGCATATAGCAACCAAATAACT
AGCTTTTGTGTATTCTGTGTTGCAAACAAATCTATGAGGTAGTTGTTCAAGGAATCCTTC
CTTGACTTATTGTATCATGGAAGTCATATGGGGGAGTCTGAAAGAGCTGTCTTCTGTAAG
CAAGGTTTGTATACACTAGTGGGGCTGGGACACCAACCCAAGCACAAAACCTAGCTATAA
CCTATCCTGGCTGCAGGATATGCTGGAACAATGGTGGCTTGGAAATTGTGGGACTGGCAA
AGCAATAGCTAGTCTAACTTGAGGCCCATTCCACAGCAGGAAGCTCATGCCCACCTCTGC
CTGGATGGCCAGGAAGCAAAATCTTGATGGCCCCAAGACCTATGGTAAACTGAACACTAC
TGGAAAAAGAAAGACTCGTGTTAATGATCTATCAAATATTTCCTAATGATATTCTGATAA
ACTCATATATTAGTCCCTGTCCTAATCATCATCACTGGGACTCCTTCCCAGCACCTGATG
GGAGCAGATAGAGATCTACATCCAAATAGTAAGTGTATCTTGGGGAACTCCACTTAAGAA
TAGAAGGAACAATTATGAGAGCCAGAGTGATCCAGAACACTAGGATCACAGAATCAACTA
AGCAGCATGCATAGGGGTTAATGGAGACTGAAGTGGCAATCACAGAGCCTGCATAGGTCT
ACACTAAGTCCTCTGTGTATATACTGTGGCTGTTTAGCTTAGGAATTTTGTTGGACTCCT
AACAATGGATAAGGAATTC

SEQ ID NO:109

Mouse T2R03 amino acid sequence

MVLTIRAILWVT[redacted]TIISLEFIIGILGNVFIALVNIIDWVKRGK[redacted]AVDKTYMALAISRT
AFLLSLITGFLVSLLDPALLGMRTMVRLLTISWMVTNHFSVWFATCLSIFYFLKIANFSN
SIFLVLKWEAKKVVSVTLVVSVIILIMNIIVINKFTDRLQVNTLQNCSTSNTLKDYGLFL
FISTGFTLTPFAVSLTMFLLLIFSLWRHLKNMCHSATGSRDVSTVAHIKGLQTVVTFLLL
YTAFVMSLLSESLNINIQHTNLLSHFLRSIGVAFPTGHSCVLILGNSKLRQASLSVILWL
RYKYKHIENWGP

SEQ ID NO:110

Mouse T2R03 nucleotide sequence

CTTTAATAGCAGGGTGTGAATATTTAAATTTTCTTTCTGCAGCAACTACTGAGGGCTTCA
GACTGCTGTATACAGGGCATGAAGCATCTGGATGAAGTTCAGCTGTGCTGCCTTTGACAA
CAATTTTTTGTGTATGTGTGGAGAACATAAACCATTTCATTAGTGAAATTTGGCTTTTGG
GTGACATTGTCTATGATAGTTCTGAAAGTGATTATGTTAAGAATCAGACACAGCCGTCTA
GAAGATTGTATTAACACATCTTTGGTAGTTCAGAAGAAATTAGATCATC**ATGGTGTTGAC**
AATAAGGGCTATTTTATGGGTAACATTGATAACTATTATAAGTCTGGAGTTTATCATAGG
AATTTTAGGAAATGTATTCATAGCTCTCGTGAACATCATAGACTGGGTTAAAAGAGGAAA
GATCTCTGCAGTGGATAAGACCTATATGGCCCTGGCCATCTCCAGGACTGCTTTTTTATT
GTCACTAATCACAGGGTTCTTGGTATCATTATTGGACCCAGCTTTATTGGGAATGAGAAC
GATGGTAAGGCTCCTTACTATTTCCTGGATGGTGACCAATCATTTCAGTGTCTGGTTTGC
AACATGCCTCAGTATCTTTTATTTTCTCAAGATAGCTAATTTCTCAAATTCTATTTTCCT
TGTTCTCAAATGGGAAGCTAAAAAAGTGGTATCAGTGACATTGGTGGTATCTGTGATAAT
CTTGATCATGAACATTATAGTCATAAACAAATTCACTGACAGACTTCAAGTAAACACACT
CCAGAACTGTAGTACAAGTAACACTTTAAAAGATTATGGGCTCTTTTTATTCATTAGCAC
TGGGTTTACACTCACCCCATTCGCTGTGTCTTTGACAATGTTTCTTCTGCTCATCTTCTC
CCTGTGGAGACATCTGAAGAATATGTGTCACAGTGCCACAGGCTCCAGAGATGTCAGCAC
AGTGGCCCACATAAAAGGCTTGCAAACTGTGGTAACCTTCCTGTTACTATATACTGCTTT
TGTTATGTCACTTCTTTCAGAGTCTTTGAATATTAACATTCAACATACAAATCTTCTTTC
TCATTTTTTACGGAGTATAGGAGTAGCTTTTCCCACAGGCCACTCCTGTGTACTGATTCT
TGGAAACAGTAAGCTGAGGCAAGCCTCTCTTTCTGTGATATTGTGGCTGAGGTATAAGTA
CAAACATATAGAGAATTGGGGCCCCTAAATCATATCAGGGATCCTTTTCCACATTCTAGA
AAAAAATCAGTTAATAAGAACAGGAATTTAGGAAGGAATCTGAAATTATGAATCTCATAG
GCCATGAACCTTCAGACAAAGGATTCATTAGAGAGATAGAGAGAGAACATTGTTATCTGT

AACTCGACAGGCA●ACTGTAGATTATGAAAATAAATGTCAGTC●TAATGGAAAGCAAA
ACATGCTATATTTTATTAATTGGTTTTGGTTTAAGGTCGGGATA

SEQ ID NO:111

Mouse T2R04 amino acid sequence

MLSALESILLSVATSEAMLGVLGNTFIVLVNYTDWVRNKKLSKINFILTGLAISRIFTIW
IITLDAYTKVFLLTMLMPSSLHECMSYIWVIINHLSVWFSTSLGIFYFLKIANFSHYIFL
WMKRRADKVFVFLIVFLIITWLASFPLAVKVIKDVKIYQSNTSWLIHLEKSELLINYVFA
NMGPISLFIVAIIACFLLTISLWRHSRQMQSIGSGFRDLNTEAHMKAMKVLIAFIILFIL
YFLGILIETLCLFLTNNKLLFIFGFTLSAMYPCCHSFILILTSRELKQDTMRALQRLKCC
ET

SEQ ID NO:112

Mouse T2R04 nucleotide sequence

CTGCAGCAGGTAAATCACACCAGATCCAGCAGAAGCCTTCTTGGAAATTGGCAGAG**ATGC**
TGAGTGCACTGGAAAGCATCCTCCTTTCTGTTGCCACTAGTGAAGCCATGCTGGGAGTTT
TAGGGAACACATTTATTGTACTTGTAAACTACACAGACTGGGTCAGGAATAAGAAACTCT
CTAAGATTAACTTTATTCTCACTGGCTTAGCAATTTCCAGGATTTTTACCATATGGATAA
TAACTTTAGATGCATATACAAAGGTTTTCCTTCTGACTATGCTTATGCCGAGCAGTCTAC
ATGAATGCATGAGTTACATATGGGTAATTATTAACCATCTGAGCGTTTGGTTTAGCACCA
GCCTCGGCATCTTTTATTTTCTGAAGATAGCAAATTTTTCCCACTACATATTTCTCTGGA
TGAAGAGAAGAGCTGATAAAGTTTTTGTCTTTCTAATTGTATTCTTAATTATAACGTGGC
TAGCTTCCTTTCCGCTAGCTGTGAAGGTCATTAAAGATGTTAAAATATATCAGAGCAACA
CATCCTGGCTGATCCACCTGGAGAAGAGTGAGTTACTTATAAACTATGTTTTTGCCAATA
TGGGGCCCATTTCCCTCTTTATTGTAGCCATAATTGCTTGTTTCTTGTTAACCATTTCCC
TTTGGAGACACAGCAGGCAGATGCAATCCATTGGATCAGGATTCAGAGATCTCAACACAG
AAGCTCACATGAAAGCCATGAAAGTTTTAATTGCATTTATCATCCTCTTTATCTTATATT
TTTTGGGTATTCTCATAGAAACATTATGCTTGTTTCTTACAAACAATAAACTTCTCTTTA
TTTTTGGCTTCACTTTGTCAGCCATGTATCCCTGTTGCCATTCCTTTATCCTAATTCTAA
CAAGCAGGGAGCTGAAGCAAGACACTATGAGGGCACTGCAGAGATTAAAATGCTGTGAGA

CTTGACAGAGAAA●AATGTTCTGGCACAGTTCAGCAGGGAATC●TGGAGCCCTTTCCA
TTCCCACTATGTTCTCACACTGTCTTTAGTTGAATTGTTAAAAGTTTTTGAAACCTTTGG
CAACTGATTGACTGCAGCTACGCCAGTGTAAGATTTTCATAGTAAGAGCAAACATTGAAA
ATAAGACTTCTCAGTCTTATTTCATTGAGTTTCTAAAGCATTGACACCCATTCACCAGAA
AAACCAAAGGGGAAGAGAGGAGTTTTCAGACATGTGTGATGAATCTTGATATTTAGGACA
TGGAATTGAGGAG~CCAGAGGGATGCTACCGTGTGTCTACAGCTTTGTTTGTTAAATAGC
TACTTTTCCTTTCCCAGTTAGTTAAAGTAGATGCTTGGAGTAGTGGTGAAAATCATGGCA
GTAGATGGGATCTGTGGGAAGTGGTTGAGGAAGCAGGCTGTTTCTGAACGAAGAGACCAG
AGGACTGATTGAACTGGTCATTGTGTATATCAAAAATAGTGATTTCAGATGAAGCCAAGT
TGTAGAGCAAAGATATCTGAGGAAGAATTC

SEQ ID NO:113

Mouse T2R05 amino acid sequence

MLSAAEGILLSIATVEAGLGVLGNTFIALVNCMDWAKNNKLSMTGFLLIGLATSRIFIVW
LLTLDAYAKLFYPSKYFSSSLIEIISYIWMTVNHLTVWFATSLSIFYFLKIANFSDCVFL
WLKRRTDKAFVFLLGCLLTSWVISFSFVVKVMKDGKVNHRNRTSEMYWEKRQFTINYVFL
NIGVISLFMMTLTACFLLIMSLWRHSRQMQSGVSGFRDLNTEAHVKAIKFLISFIILFVL
YFIGVSIEIICIFIPENKLLFIFGFTTASIYPCCHSFILILSNSQLKQAFVKVLQGLKFF

SEQ ID NO:114

Mouse T2R05 nucleotide sequence

ATGCTGAGTGCGGCAGAAGGCATCCTCCTTTCCATTGCAACTGTTGAAGCTGGGCTGGGA
GTTTTAGGGAACACATTTATTGCACTGGTAAACTGCATGGACTGGGCCAAGAACAATAAG
CTTTCTATGACTGGCTTCCTTCTCATCGGCTTAGCAACTTCCAGGATTTTTATTGTGTGG
CTATTAACTTTAGATGCATATGCAAAGCTATTCTATCCAAGTAAGTATTTTTCTAGTAGT
CTGATTGAAATCATCTCTTATATATGGATGACTGTGAATCACCTGACTGTCTGGTTTGCC
ACCAGCCTAAGCATCTTCTATTTCCTGAAGATAGCCAATTTTTCCGACTGTGTATTTCTC
TGGTTGAAGAGGAGAACTGATAAAGCTTTTGTTTTTCTCTTGGGGTGTTTGCTAACTTCA
TGGGTAATCTCCTTCTCATTTGTTGTGAAGGTGATGAAGGACGGTAAAGTGAATCATAGA
AACAGGACCTCGGAGATGTACTGGGAGAAAAGGCAATTCACTATTAACTACGTTTTCCTC

AATATTGGAGTCA[redacted]TCTCTCTTTATGATGACCTTAACTGCATG[redacted]TCTTGTTAATTATG
TCACTTTGGAGACACAGCAGGCAGATGCAGTCTGGTGTTTCAGGATTCAGAGACCTCAAC
ACAGAAGCTCATGTGAAAGCCATAAAATTTTTAATTTCATTTATCATCCTTTTCGTCTTG
TATTTTATAGGTGTTTCAATAGAAATTATCTGCATATTTATACCAGAAAACAAACTGCTA
TTTATTTTTGGTTTCACAACTGCATCCATATATCCTTGCTGTCACTCATTTATTCTAATT
CTATCTAACAGCCAGCTAAAGCAAGCCTTTGTAAAGGTACTGCAAGGATTAAAGTTCTTT
TAG

SEQ ID NO:115

Mouse T2R06 amino acid sequence

MLTVAEGILLCFVTSGSVLGVLGNGFILHANYINCVRKKFSTAGFILTGLAICRIFVICI
IISDGYLKLFSPHMVASDAHIIVISYIWVIINHTSIWFATSLNLFYLLKIANFSHYIFFC
LKRRINTVFIFLLGCLFISWSIAFPQTVKIFNVKKQHRNVSWQVYLYKNEFIVSHILLNL
GVIFFFMVAIITCFLLIISLWKHNRKMQLYASRFKSLNTEVHVKVMKVLISFIILLILHF
IGILIETLSFLKYENKLLLILGLIISCMYPCCHSFILILANSQLKQASLKALKQLKCHKK
DKDVRVTW

SEQ ID NO:116

Mouse T2R06 nucleotide sequence

TATAGTTGCAGCAGAAGCAACGTTAGGGATCTGTAGAGATGCTGACTGTAGCAGAAGGAA
TCCTCCTTTGTTTTGTAACTAGTGGTTCAGTCCTGGGAGTTCTAGGAAATGGATTTATCC
TGCATGCAAACTACATTAACTGTGTCAGAAAGAAGTTCTCCACAGCTGGCTTTATTCTCA
CAGGCTTGGCTATTTGCAGAATCTTTGTCATATGTATAATAATCTCTGATGGATATTTAA
AATTGTTTTCTCCACATATGGTTGCCTCTGATGCCCACATTATAGTGATTTCTTACATAT
GGGTAATTATCAATCATACAAGTATATGGTTTGCCACCAGCCTCAACCTCTTCTATCTCC
TGAAGATAGCAAATTTTTCTCACTACATCTTCTTCTGCTTGAAGAGAAGAATCAATACAG
TATTTATCTTTCTCCTGGGATGCTTATTTATATCATGGTCAATTGCTTTCCCACAAACAG
TGAAGATATTTAATGTTAAAAAGCAGCACAGAAaTGTTTCCTGGCAGGTTTACCTCTATA
AGAATGAGTTCATtGTAAGCCACATTCTTCTCAACCTGGGAGTTATATTCTTCTTTATGG
TGGCTATCATTACATGCTTCCTATTAATTATTTCACTTTGGAAACATAACAGAAAGATGC

AGTTGTATGCCT[redacted]AGATTCAAAAGCCTTAACACAGAAGTACA[redacted]TGAAAGTCATGAAAG
TTTTAATTTCTTTTATTATCCTGTTAATCTTGCATTTCATAGGGATTTTGATAGAAACAT
TGAGCTTTTTAAAATATGAAAATAAACTGCTACTTATTTTGGGTTTGATAATTTCATGCA
TGTATCCTTGCTGTCATTCATTTATCCTAATTCTAGCAAACAGTCAGCTGAAGCAGGCTT
CTTTGAAGGCACTGAAGCAATTAAAATGCCATAAGAAAGACAAGGACGTCAgAGTGACAT
GGTAGACTTATGGAGAAATGAATGGTCACAAGAAATAGCCTGGTGTGGAGATGTTGATAT
CTCTAAAGACCGTTTCACTTCCAAATTCTTGCAATTATTTAAAAAAAAAAGTCTTGCTGA
TATCATGGAATCATGGGAAATGTTGCAATTGTGTTTTGGGGACAGGGTGACCAGTGAAGG
TATGGTTAAGCAGCGAAACACTCATACAGCTCGTTCGTTCTTTTTGTATTTTATTTTGTG
TTGGTGGCCTTCCAAGACATGATTTCTCTATGTAAGTTTTGG

SEQ ID NO:117

Mouse T2R07 amino acid sequence

MLNSAEGILLCVVTSEAVLGVLGDTYIALFNCMDYAKNKKLSKIGFILIGLAISRIGVVW
IIILQGYIQVFFPHMLTSGNITEYITYIWVFLNHLSVWFVTNLNILYFLKIANFSNSVFL
WLKRRVNAVFIFLSGCLLTSWLLCFPQMTKILQNSKMHQRNTSWVHQRKNYFLINQSVTN
LGIFFFIIVSLITCFLLIVFLWRHVRQMHSDVSGFRDHSTKVHVKAMKFLISFMVFFILH
FVGLSIEVLCFILPQNKLLFITGLTATCLYPCGHSIIVILGNKQLKQASLKALQQLKCCE
TKGNFRVK

SEQ ID NO:118

Mouse T2R07 nucleotide sequence

TTCATAATGAAGAGGAGGCAGGGCAATGTTGGTTTCTGTTGTCTGACCAGTGTATTTGAC
AGTGATACTACACATTTGATTGCTAAATGCAAATAGTTCCAAAGGAACAAGTAAATTTTA
TGAAATAGAAGCTTCTATTTGCTTATTAACAAACTGCAAGCAAACATTAGTCTGCACACA
TTTTATAGACAAGCTAAATCTTCAAAAGCAATAAAAAAAGAGCACCCATAAAGTTCTGACT
CTATCACATGACAATAGGCTTGAAAAGATTGTCTATGTAGATAAAGAAGATGGCATAACT
TCTCCATCAAGAAGCCAGTATATGGGACATTCTCCAGCAGATAATTTACAATAGATGCAG
CAGAAGTAACCTTAGAGATCTGTAAAG**ATGCTGAATTCAGCAGAAGGCATCCTCCTTTGT**
GTTGTCACTAGTGAGGCTGTGCTCGGAGTTTTAGGGGACACATATATTGCACTTTTTAAC

TGCATGGACTATC[redacted]AAGAACAAGAAGCTCTCTAAGATCGGTTT[redacted]TTCTCATTGGCTTG
GCGATTTCCAGAATTGGTGTTGTATGGATAATAATTTTACAAGGGTATATACAAGTATTT
TTTCCACACATGCTTACCTCTGGAAACATAACTGAATATATTACTTACATATGGGTATTT
CTCAATCACTTAAGTGTCTGGTTTGTCACCAACCTCAACATCCTCTACTTTCTAAAGATA
GCTAATTTTTCCAACTCTGTATTTCTCTGGCTGAAAAGGAGAGTCAATGCAGTTTTTATC
TTTCTGTCAGGATGCTTACTTACCTCATGGTTACTATGTTTTCCACAAATGACAAAGATA
CTTCAAAATAGTAAAATGCACCAGAGAAACACATCTTGGGTCCACCAGCGGAAAAATTAC
TTTCTTATTAACCAAAGTGTGACCAATCTGGGAATCTTTTTCTTCATTATTGTATCCCTG
ATTACCTGCTTTCTGTTGATTGTTTTCCTCTGGAGACATGTCAGACAAATGCACTCAGAT
GTTTCAGGATTCAGAGACCACAGCACAAAAGTACATGTGAAAGCTATGAAATTTCTAATA
TCTTTTATGGTCTTCTTTATTCTGCATTTTGTAGGCCTTTCCATAGAAGTGCTATGCTTT
ATTCTGCCACAAAATAAACTGCTCTTTATAACTGGTTTGACAGCCACATGCCTCTATCCC
TGCGGTCACTCAATCATCGTAATTTTAGGAAATAAGCAGTTAAAGCAAGCCTCTTTGAAG
GCACTGCAGCAACTAAAATGCTGTGAGACAAAAGGAAATTTCAGAGTCAAATAAATGGGT
TTGCAAATAAATAGCTGCCTTGTTCTTcCACTGGTTTTTACCCTGTTAGTTGATGTTATG
AAAAGTTCCTGCTATGGTTGATGACATCTCAAGGAATCTATTTTTCTGGTGGCATGTTAA
GTCCACGTGAAGCCTCACTTCATACTGTGACTTGACTATGCAAATTCTTTCCACAAAATA
ACCAGATAACATTCAGCCTGGAGATAAATTCATTTAAAGGCTTTTATGGTGAGGATAAAC
AAAAAAAAAAAATCATTTTTCTGTGATTCACTGTAACTCCCAGGATGAGTAAAAGAAAAC
AAGACAAATGGTTGTGATCAGCCTTTGTGTGTCTAGACAGAGCTAGGGACCAGATGTTGA
TGCTTGTGTGTGGTTTTGAGTTCTTTAAGAAGTTATTGCCTCTCTGCCATTCGGTATTCC
TCAGGTGAGAATTC

SEQ ID NO:119

Mouse T2R08 amino acid sequence

MLWELYVFVFAASVFLNFVGIIANLFIIVIIIKTWVNSRRIASPDRILFSLAITRFLTLG
LFLLNSVYIATNTGRSVYFSTFFLLCWKFLDANSLWLVTILNSLYCVKITNFQHPVFLLL
KRTISMKTTSLLLACLLISALTTLLYYMLSQISRFPEHIIGRNDTSFDLSDGILTLVASL
VLNSLLQFMLNVTFASLLIHSLRRHIQKMQRNRTSFWNPQTEAHMGAMRLMICFLVLYIP
YSIATLLYLPSYMRKNLRAQAICMIITAAYPPGHSVLLIITHHKLKAKAKKIFCFYK

SEQ ID NO:120

Mouse T2R08 nucleotide sequence

AAGCTTGTTTGTAATTAGGCATTCCTAAGAAAATAAGAACAGGAGTGAAGAAATAGTAAT
TTAATCCTTGAAAGATTTGCATCTCAGTAAAAGCAGCTGCCTCTTAGACCAGAAATGGTG
TTTGCCATGCTGGAAAATAAAAAGGAGACCTCTTTCCAGGCTGCATCCTGTGTCTGCTTA
CTTATTTCAGTTTGTTTTCATCGGCACCAAACGAGGAAAG**ATGCTCTGGGAACTGTATGT**
ATTTGTGTTTGCTGCCTCGGTTTTTTTAAATTTTGTAGGAATCATTGCAAATCTATTTAT
TATAGTGATAATTATTAAGACTTGGGTCAACAGTCGCAGAATTGCCTCTCCGGATAGGAT
CCTGTTCAGCTTGGCCATCACTAGATTCCTGACTTTGGGGTTGTTTCTACTGAACAGTGT
CTACATTGCTACAAATACTGGAAGGTCAGTCTACTTTTCCACATTTTTTCTATTGTGTTG
GAAGTTTCTGGATGCAAACAGTCTCTGGTTAGTGACCATTCTGAACAGCTTGTATTGTGT
GAAGATTACTAATTTTCAACACCCAGTGTTTCTCCTGTTGAAACGGACTATCTCTATGAA
GACCACCAGCCTGCTGTTGGCCTGTCTTCTGATTTCAGCCCTCACCACTCTCCTATATTA
TATGCTCTCACAGATATCACGTTTTCCTGAACACATAATTGGGAGAAATGACACGTCATT
TGACCTCAGTGATGGTATCTTGACGTTAGTAGCCTCTTTGGTCCTGAACTCACTTCTACA
GTTTATGCTCAATGTGACTTTTGCTTCCTTGTTAATACATTCCTTGAGAAGACATATACA
GAAGATGCAGAGAAACAGGACCAGCTTTTGGAATCCCCAGACGGAGGCTCACATGGGTGC
TATGAGGCTGATGATCTGTTTCCTCGTGCTCTACATTCCATATTCAATTGCTACCCTGCT
CTATCTTCCTTCCTATATGAGGAAGAATCTGAGAGCCCAGGCCATTTGCATGATTATTAC
TGCTGCTTACCCTCCAGGACATTCTGTCCTCCTCATTATCACACATCATAAACTGAAAGC
TAAAGCAAAGAAGATTTTCTGTTTCTACAAGTAGCAGAATTTCATTAGTAGTTAACAGCA
TCAATTCATGGTTTGGTTGCATTAGAAATGTCTCAGTGATCTAAGGACTTAATTTTGTGA
TCTTGTATCTGGCATCCTGACCCTGAGACTAAGTGCTTATATTTTGGTCAATACAGCATC
TTTTGGCTAATATTTTAAAGTAAATCACATTCCATAAGAAATTGTTTAAGGGATTTACGT
ATTTTTCATGGCTATCACATTCCTAGACAATGGAAATCACCATACTGTTTCGCTAGCTAC
TGAAGTACCAGGGGAAAGTCCATGAATGAAGGCCACATTGTGATGTTCTTGGTTAGCACA
GATTAGAGAATTTGGCCTCAACTGAGCAAGATATC

SEQ ID NO:121

Mouse T2R09 amino acid sequence

MEHLLKRTFDITE[redacted]LLIILFIELIIGLIGNGFTALVHCMDWVK[redacted]KMSLVNKILTALAT
SRIFLLWFMLVGFPISSLYPYLVTTRLMIQFTSTLWTIANHISVWFATCLSVFYFLKIAN
FSNSPFLYLKRRVEKVVSVTLLVSLVLLFLNILLLNLEINMCINEYHQINISYIFISYYH
LSCQIQVLGSHIIFLSVPVVLSLSTFLLLIFSLWTLHKRMQQHVQGGRDARTTAHFKALQ
AVIAFLLLYSIFILSLLLQFWIHGLRKKPPFIAFCQVVDTAFPSFHSYVLILRDRKLRHA
SLSVLSWLKCRPNYVK

SEQ ID NO:122

Mouse T2R09 nucleotide sequence

GAATTCAGAAATCATCAAAAAATCTTCAAAACTACATGTTTAAAATAGCACTTCAAATGA
ATACATTTGCAAATCTTTACAACTAATACATAAA**ATGGAGCATCTTTTGAAGAGAACATT**
TGATATCACCGAGAACATACTTCTAATTATTTTATTCATTGAATTAATAATTGGACTTAT
AGGAAACGGATTCACAGCCTTGGTGCACTGCATGGACTGGGTTAAGAGAAAAAAAATGTC
ATTAGTTAATAAAATCCTCACCGCTTTGGCAACTTCTAGAATTTTCCTGCTCTGGTTCAT
GCTAGTAGGTTTTCCAATTAGCTCACTGTACCCATATTTAGTTACTACTAGACTGATGAT
ACAGTTCACTAGTACTCTATGGACTATAGCTAACCATATTAGTGTCTGGTTTGCTACATG
CCTCAGTGTCTTTTATTTTCTCAAGATAGCCAATTTTTCTAATTCTCCTTTTCTCTATCT
AAAGAGGAGAGTTGAAAAAGTAGTTTCAGTTACATTACTGGTGTCTCTGGTCCTCTTGTT
TTTAAATATTTTACTACTTAATTTGGAAATTAACATGTGTATAAATGAATATCATCAAAT
AAACATATCATACATCTTCATTTCTTATTACCATTTAAGTTGTCAAATTCAGGTGTTAGG
AAGTCACATTATTTTCCTGTCTGTCCCCGTTGTTTTGTCCCTGTCAACTTTTCTCCTGCT
CATCTTCTCCCTGTGGACACTTCACAAGAGGATGCAGCAGCATGTTCAGGGAGGCAGAGA
TGCCAGAACCACGGCCCACTTCAAAGCCTTGCAAGCAGTGATTGCCTTTCTCCTACTATA
CTCCATTTTTATCCTGTCACTGTTACTACAATTTTGGATCCATGGATTAAGGAAGAAACC
TCCTTTCATTGCATTTTGTCAGGTTGTAGATACAGCTTTTCCTTCATTCCATTCATATGT
CTTGATTCTGAGAGACAGGAAGCTGAGACACGCCTCTCTCTCTGTGTTGTCGTGGCTGAA
ATGCAGGCCAAATTATGTGAAATAATATTTCTTTGTATTTTCATTTTCAATTTTAAAATA
TTCTTAGAATTTGACTGCATGTATTTCATCTTTTATTTGAAACAACCACTAATTAAAGCT
ATTACTAATTTAGCAAGTCGTATACAAGGTTATTTTTTAATACACATATCAAAAACTGAC
ATGTTTATGTTCTACAAAAACCTGAATATATCAAAATTATATAAATTTTGTATCAACGAT
TAACAATGGAGTTTTTTTATTTATGACCTGTCACGGGACTCCGGTGGAGTCAGCTTGTCA
GATGAAAGTCTGAAAGCTT

SEQ ID NO:123

Mouse T2R10 amino acid sequence

MFSQIISTSDIFTFTIILFVELVIGILGNGFIALVNIMDWTKRRSISSADQILTALAITR
FLYVWFMIICILLFMLCPHLLTRSEIVTSIGIIWIVNNHFSVWLATCLGVFYFLKIANFS
NSLFLYLKWRVKKVVLMIIQVSMIFLILNLLSLSMYDQFSIDVYEGNTSYNLGDSTPFPT
ISLFINSSKVFVITNSSHIFLPINSLFMLIPFTVSLVAFLMLIFSLWKHHKKMQVNAKPP
RDASTMAHIKALQTGFSFLLLYAVYLLFIVIGMLSLRLIGGKLILLFDHISGIGFPISHS
FVLILGNNKLRQASLSVLHCLRCRSKDMDTMGP

SEQ ID NO:124

Mouse T2R10 nucleotide sequence

GAATTCAACATCTTATTCAACTTCAGAAAACTGGATATTAGACACAGTGTCTGGATGAAG
CAGAGGTGATCTCTTTGGGAAAAAAAGCCAAGTAGTCATAAAGAATTTATGAAACAATTC
CTGGGATTGTTTATATTTGTTACAAACAAATTTATATGTTTGTTAGTCAGTAATGTATAA
GTGGGATTTTAAAGCATGATTATCTTGAATTTTTAACAAAAAACATGTAGTGCTTTTTAA
ATGTAGCAGAAACATTAAAAATTGAAGC**ATGTTCTCACAGATAATAAGCACCAGTGATAT**
TTTTACTTTTACAATAATATTATTTGTGGAATTAGTAATAGGAATTTTAGGAAATGGATT
CATAGCACTAGTGAATATCATGGACTGGACCAAGAGAAGAAGCATTTCATCAGCGGATCA
GATTCTCACTGCTTTGGCCATTACCAGATTTCTCTATGTGTGGTTTATGATCATTTGTAT
ATTGTTATTCATGCTGTGCCCACATTTGCTTACAAGATCAGAAATAGTAACATCAATTGG
TATTATTTGGATAGTGAATAACCATTTCAGCGTTTGGCTTGCCACATGCCTCGGTGTCTT
TTATTTTCTGAAGATAGCCAATTTTTCTAACTCTTTGTTTCTTTACCTAAAGTGGAGAGT
TAAAAAAGTAGTTTTAATGATAATACAGGTATCAATGATTTTCTTGATTTTAAACCTGTT
ATCTCTAAGCATGTATGATCAGTTCTCAATTGATGTTTATGAAGGAAATACATCTTATAA
TTTAGGGGATTCAACCCCATTTCCCACAATTTCCTTATTCATCAATTCATCAAAAGTTTT
CGTAATCACCAACTCATCCCATATTTTCTTACCCATCAACTCCCTGTTCATGCTCATACC
CTTCACAGTGTCCCTGGTAGCCTTTCTCATGCTCATCTTCTCACTGTGGAAGCATCACAA
AAAGATGCAGGTCAATGCCAAACCACCTAGAGATGCCAGCACCATGGCCCACATTAAAGC
CTTGCAAACAGGGTTCTCCTTCCTGCTGCTGTATGCAGTATACTTACTTTTTATTGTCAT

AGGAATGTTGAG●TTAGGTTGATAGGAGGAAAATTAATACTT●ATTTGACCACATTTC
TGGAATAGGTTTTCCTATAAGCCACTCATTTGTGCTGATTCTGGGAAATAACAAGCTGAG
ACAAGCCAGTCTTTCAGTGTTGCATTGTCTGAGGTGCCGATCCAAAGATATGGACACCAT
GGGTCCATAAAAAATTTCAGAGGTCATTGGGAAACATTTTGAGATCTTATAGGGGAAAAA
GAAAATGTGGGGCTTCAAAGCTGGTAGGAGTAATATAGAGAAGGATAGGAG

SEQ ID NO:125

Mouse T2R11 amino acid sequence

MEHPLRRTFDFSQSILLTILFIELIIGLIRNGLMVLVHCIDWVKRKKFHLLIKSSPLWQT
SRICLLWFMLIHLLITLLYADLASTRTMMQFASNPWTISNHISIWLATCLGVFYFLKIAN
FSNSTFLYLKWRVQFLLLNILLVKFEINMWINEYHQINIPYSFISYYQXCQIQVLSLHII
FLSVPFILSLSTFLLLIFSLWTLHQRMQQHVQGYRDASTMAHFKALQAVIAFLLIHSIFI
LSLLLQLWKHELRKKPPFVVFCQVAYIAFPSSHSYVFILGDRKLRQACLSVLWRLKCRPN
YVG

SEQ ID NO:126

Mouse T2R11 nucleotide sequence

AATAATGTATGTGGAAGAGTTAAGTATAAATGTTGTATGAGAATGAACTCAGAAATCATC
AAAAATCTTTAAAACTGCATGTTAAAAATCACACTTCAAATGAATATATTTGTAATTCTT
TAGAACTAATAAATAAAATGGAGCATCCTTTGAGGAGAACATTTGATTTCTCCCAGAGCA
TACTTCTAACCATTTTATTCATTGAATTAATAATTGGACTTATAAGAAATGGATTAATGG
TATTGGTGCACTGCATAGATTGGGTTAAGAGAAAAAAATTTCATTTGTTAATCAAATCCT
CACCACTTTGGCAAACTTCCAGAATTTGTCTGCTCTGGTTCATGCTAATACATCTCCTGA
TTACTTTATTGTATGCAGATTTAGCTAGTACTAGAACGATGATGCAATTCGCTAGCAATC
CATGGACTATATCTAACCATATCAGCATCTGGCTTGCTACATGCCTTGGTGTCTTTTATT
TTCTCAAGATAGCCAATTTTTCTAACTCTACTTTTCTCTATCTAAAATGGCGAGTTCAGT
TCCTCTTGTTAAATATTTTACTGGTTAAATTTGAGATTAACATGTGGATAAATGAATATC
ATCAAATAAACATACCATACAGCTTCATTTCTTATTACCAAATTGTCAAATACAGGTGTT
AAGTCTTCACATTATTTTCCTGTCTGTCCCCTTTATTTTGTCCCTGTCAACTTTTCTCCT
GCTCATCTTCTCCCTGTGGACACTTCACCAGAGGATGCAGCAGCATGTTCAAGGATACAG

AGATGCCAGCAC[redacted]TGGCCCACTTCAAAGCCTTGCAAGCAGTGA[redacted]GCCTTTCTCTTAAT
ACACTCCATTTTTATCCTGTCACTGTTACTACAACTTTGGAAACATGAATTAAGGAAGAA
ACCTCCTTTTGTTGTATTTTGTCAGGTTGCATATATAGCTTTTCCTTCATCCCATTCATA
TGTCTTCATTCTGGGAGACAGAAAGCTGAGACAGGCTTGTCTCTCTGTGTTGTGGAGGCT
GAAATGCAGGCCAAATTATGTGGGATAAAATCTCTTTGTGCTTTCATTTCCAATTCTTAA
ATATTCTTTGATTTTGACTGCATAAATT

SEQ ID NO:127

Mouse T2R12 amino acid sequence

GAIVNVDFLIGNVGNGFIVVANIMDLVKRRKLSSVDQLLTALAVSRITLLWYLYIMKRTF
LVDPNIGAIMQSTRLTNVIWIISNHFSIWLATTLSIFYFLKIANFSNSIFCYLRWRFEKV
ILMALLVSLVLLFIDILVTNMYINIWTDEF

SEQ ID NO:128

Mouse T2R12 nucleotide sequence

TTTTCAGCAGTGACTTTGGGAAGCAGAACGTCCTCTTAGAGACAGTGGGTGCTGCTATCC
TAGTTAATGTGGAGCAATAGTTAATGTGGATTTCCTAATTGGAAATGTTGGGAATGGATT
CATTGTTGTGGCAAACATAATGGACTTGGTCAAGAGAAGAAAGCTTTCTTCAGTGGATCA
GCTGCTCACTGCACTGGCCGTCTCCAGAATCACTTTGCTGTGGTACCTGTACATAATGAA
ACGAACATTTTTAGTGGATCCAAACATTGGTGCAATTATGCAATCAACAAGACTGACTAA
TGTTATCTGGATAATTTCTAACCATTTTAGTATATGGCTGGCCACCACCCTCAGCATCTT
TTATTTTCTCAAGATAGCAAATTTTTCTAACTCTATTTTCTGTTACCTGAGGTGGAGATT
TGAAAAGGTGATTTTGATGGCATTGCTGGTGTCCCTGGTCCTCTTGTTTATAGATATTTT
AGTAACAAACATGTACATTAATATTTGGACTGATGAATTC

SEQ ID NO:129

Mouse T2R13 amino acid sequence

MVAVLQSTLPIIF●EFIMGTLGNGFIFLIVCIDWVQRRKISLV●IRTALAISRIALIW
LIFLDWWVSVHYPALHETGKMLSTYLISWTVINHCNFWLTANLSILYFLKIANFSNIIFL
YLKFRSKNVVLVTLLVSLFFLFLNTVIIKIFSDVCFDSVQRNVSQIFIMYNHEQICKFLS
FTNPMFTFIPFVMSTVMFSLLIFSLWRHLKNMQHTAKGCRDISTTVHIRALQTIIVSVVL
YTIFFLSFFVKVWSFVSPERYLIFLFVWALGNAVFSAHPFVMILVNRRLRLASLSLIFWL
WYRFKNIEV

SEQ ID NO:130

Mouse T2R13 nucleotide sequence

AAGCTTGTTTGTGTTTGGATGAATTCTATTTATGTCTATCAATTTAAGATTTTCATATGA
ATCATTAAGAAATCTTGATAGTTGTTTGTGAGATATCACTTCTGCAATTTTTAAATGAAA
TTACACTCATATTTTGAAGGAACAATATGTTTTAAAGGAATATATTAACAAATCTTCAGC
AGTTACCTCAGAAGTTTGGGTATTGTTTTACAGAAA**ATGGTGGCAGTTCTACAGAGCACA**
CTTCCAATAATTTTCAGTATGGAATTCATAATGGGAACCTTAGGAAATGGATTCATTTTT
CTGATAGTCTGCATAGACTGGGTCCAAAGAAGAAAAATCTCTTTAGTGGATCAAATCCGC
ACTGCTCTGGCAATTAGCAGAATCGCTCTAATTTGGTTGATATTCCTAGATTGGTGGGTG
TCTGTTCATTACCCAGCATTACATGAAACTGGTAAGATGTTATCAACATATTTGATTTCC
TGGACGGTGATCAATCATTGTAACTTTTGGCTTACTGCAAACTTGAGCATCCTTTATTTT
CTCAAGATAGCCAACTTTTCTAACATTATTTTTCTTTATCTAAAGTTTAGATCTAAAAAT
GTGGTATTAGTGACCCTGTTAGTGTCTCTATTTTTCTTGTTCTTAAATACTGTAATTATA
AAAATATTTTCTGATGTGTGTTTTGATAGTGTTCAAAGAAATGTGTCTCAAATTTTCATA
ATGTATAACCATGAACAAATTTGTAAATTTCTTTCCTTTACTAACCCTATGTTCACATTC
ATACCTTTTGTTATGTCCACGGTAATGTTTTCTTTGCTCATCTTCTCCCTGTGGAGACAT
CTGAAGAATATGCAGCACACCGCCAAAGGATGCAGAGACATCAGCACCACAGTGCACATC
AGAGCCCTGCAAACCATCATTGTGTCTGTAGTGCTATACACTATTTTTTTTCTATCATTT
TTTGTTAAAGTTTGGAGTTTTGTGTCACCAGAGAGATACCTGATCTTTTTGTTTGTCTGG
GCTCTGGGAAATGCTGTTTTTTCTGCTCACCCATTTGTCATGATTTTGGTAAACAGAAGA
TTGAGATTGGCTTCTCTCTCTCTGATTTTTTGGCTCTGGTACAGGTTTAAAAATATAGAA
GTATAGGGTCCAAAGACCACCAAGGAATCATTTTCCTTATCCTAAAGAAAAATCAGGAG

SEQ ID NO:131

Mouse T2R14 amino acid sequence

MLSTMEGVLLSVSTSEAVLGIVGNTFIALVNCMDYNRNKKLSNIGFILTGLAISRICLVL
ILITEAYIKIFYPQLLSPVNIIELISYLWIIICQLNVWFATSLSIFYFLKIANFSHYIFV
WLKRRIDLVFFFLIGCLLISWLFSFPVVAKMVKDNKMLYINTSWQIHMKKSELIINYVFT
NGGVFLFFMIMLIVCFLLIISLWRHRRQMESNKLGFRDLNTEVHVRTIKVLLSFIILFIL
HFMGITINVICLLIPESNLLFMFGLTTAFIYPGCHSLILILANSRLKQCSVMILQLLKCC
ENGKELRDT

SEQ ID NO:132

Mouse T2R14 nucleotide sequence

CTGCAGGTATATACCTACCCTGAAGGCTTCATCTAGAGTAAACAAAGTAGTCTGTATAGT
CTGCCATTCCTCAGATTCTCCTCAACTTCCCACCCTCCAGTGACCTTTCTCCTTTTCTAC
AGTCAAACTATGGACCTCACAACCTGACACTTCTTCAGATGCAAAATATTCTCACAGAGA
CAAGTAAAACATACAAAACAAATACTTTAATTTGCCTATTAACAAATGGCAAGAAAAGAT
TCAGGCTTGAACATCCTGTAGACAAGCTAAGGACAGGAGCAACTGAAGGGATCTCCATGA
AGACCTTTCAGATTTCTACCAAAAGTAATTTTTAACTATATTTAAGTCTTTAAAGAAAGA
AAGTAAAGCCACTCTTTTATTGAACAGCAATAGATTGGAATCTTAAACAACTGCAACAGA
AGCCATTTTAAAGATCAACAAAGATGCTGAGCACA**ATGGAAGGTGTCCTCCTTTCAGTTT**
CAACTAGTGAGGCTGTGCTGGGCATTGTAGGGAACACATTCATTGCACTTGTAAACTGTA
TGGACTATAACAGGAACAAGAAGCTCTCTAATATTGGCTTTATTCTCACTGGCTTGGCAA
TTTCCAGAATTTGCCTTGTGTTGATCTTAATCACAGAGGCATACATAAAAATATTCTATC
CACAGTTGCTGTCTCCTGTCAACATAATTGAGCTCATCAGTTATCTATGGATAATTATCT
GTCAATTGAATGTCTGGTTTGCCACTAGTCTCAGTATTTTTTATTTCCTGAAGATAGCAA
ATTTTTCCCACTACATATTTGTCTGGTTAAAAAGAAGAATTGATTTAGTTTTTTTCTTCC
TGATAGGGTGCTTGCTTATCTCATGGCTATTTTCTTTCCCAGTTGTTGCGAAGATGGTTA
AAGATAATAAAATGCTGTATATAAACACATCTTGGCAGATCCACATGAAGAAAAGTGAGT
TAATCATTAACTATGTTTTCACCAATGGGGGAGTATTTTTATTTTTTATGATAATGTTAA
TTGTATGTTTCCTGTTAATCATTTCACTTTGGAGACATCGCAGGCAGATGGAATCAAATA
AATTAGGATTCAGAGATCTCAACACAGAAGTTCATGTGAGAACAATAAAAGTTTTATTGT
CTTTTATTATCCTTTTTATATTGCATTTCATGGGTATTACCATAAATGTAATTTGTCTGT
TAATCCCAGAAAGCAACTTGTTATTCATGTTTGGTTTGACAACTGCATTCATCTATCCCG

GCTGCCACTCAC[redacted]ATCCTAATTCTAGCAAACAGTCGGCTGAA[redacted]AGTGCTCTGTAATGA
TACTGCAACTATTAAAGTGCTGTGAGAATGGTAAAGAACTCAGAGACACATGACAGTCTG
GAACACATGCAATCTGGAATTGTCAGTGGAAAAAGTTACTGAAGATCTTTTCACTTGCAC
TATGCTCTTTTATTGATTTGGCATCATTATCAAACACTGTTGGAGCCTTGTGAACTCTTG
TTCAGAGTCTTCTGCCTCTCAAGGAATCACACTCC

SEQ ID NO:133

Mouse T2R15 amino acid sequence

MCAVLRSILTIIFILEFFIGNLGNGFIALVQCMDLRKRRTFPSADHFLTALAISRLALIW
VLFLDSFLFIQSPLLMTRNTLRLIQTAWNISNHFSIWFATSLSIFYLFKIAIFSNYLFFY
LKRRVKRVVLVILLLSMILLFFNIFLEIKHIDVWIYGTKRNITNGLSSNSFSEFSRLILI
PSLMFTLVPFGVSLIAFLLLIFSLMKHVRKMQYYTKGCKDVRTMAHTTALQTVVAFLLLY
TTFFLSLVVEVSTLEMDESLMLLFAKVTIMIFPSIHSCIFILKHNKLRQDLLSVLKWLQY
WCKREKTLDS

SEQ ID NO:134

Mouse T2R15 nucleotide sequence

AATAATAGATTTTTTAATATTCAGAATTTTTAAGTAATGTAGTATTGTTAGCAGCATAGC
TTATAGGAAAAGTTCCAAGTAATTTTGATTTTGTAATTCTGATTCCCCCAAATCAAGTAT
CAAGTTTACCTGCACAGACAAGGGAAGAAGTGGCAAAATGTGCAAATGAGAGCAACTTTA
TTTGACTGTCAGTACGTTGAAATTCAGTGTTTCCTTAATCAGTTATGGATTGACATTTAT
GTGCACAGAACCTGGAAGAATTTCAGCCAAGCTGGAGGTAAAAATCCAAAATTCTGATGA
TAAAACCAAAAGTAAATCACAGGTAAATCTTCTTTATTTTTCTTTTTTAATACTGTATAT
GGACATTTTTTAATACAGCATATTTTTTTTTTGAAATTTAGAAAAAAACCACTAAGAAAT
ATTCACCAATGGAATAGACTTTAAAGTCACTTAGAG**AATGTGTGCTGTTCTACGTAGCAT**
ACTGACAATCATTTTCATTTTGGAGTTCTTCATTGGAAATCTGGGGAATGGATTCATAGC
TCTGGTACAATGCATGGACTTACGAAAGAGAAGAACGTTCCCTTCAGCAGATCATTTCCT
CACTGCTCTGGCCATCTCCAGGCTTGCTCTGATATGGGTTTTATTTCTAGATTCATTTCT
GTTTATACAATCCCCATTACTGATGACTAGAAATACATTAAGACTGATTCAGACTGCCTG
GAATATAAGCAATCATTTCAGTATATGGTTTGCTACCAGCCTCAGCATCTTTTATCTCTT

CAAGATAGCCATT[redacted]TTCTAACTATCTTTTCTTCTACCTGAAGC[redacted]AGAGTTAAAAGGGT
GGTTTTGGTGATACTGCTGCTATCCATGATCCTTTTGTTTTTTAATATATTTTTAGAAAT
CAAACATATTGATGTCTGGATCTATGGAACCAAAAGAAACATAACTAATGGTTTGAGTTC
AAACAGTTTTTCAGAGTTTTCCAGGCTTATTTTAATTCCAAGTTTAATGTTCACATTAGT
ACCCTTTGGTGTATCCTTGATAGCTTTCCTCCTCCTAATCTTTTCCCTTATGAAACATGT
AAGGAAGATGCAGTACTACACCAAAGGATGCAAAGATGTCAGAACCATGGCCCACACCAC
AGCCCTGCAGACTGTGGTTGCCTTCCTCCTATTATATACTACTTTCTTTCTGTCTCTAGT
TGTGGAAGTTTCAACACTTGAAATGGATGAAAGTCTGATGCTTCTGTTTGCAAAAGTTAC
TATAATGATTTTTCCTTCCATCCACTCCTGTATTTTCATTTTGAAACATAATAAGTTGAG
ACAGGACTTGCTTTCAGTACTGAAGTGGCTACAGTATTGGTGCAAGCGTGAGAAAACCTT
GGATTCATAGACCATTGTATGCATCACCTTGAATATTCTAGAGGGGTGTAGGTTCATATG
AAAGTATTGAATTTTTAAATTTGAGCCTTTTGTATATTTTCT

SEQ ID NO:135

Mouse T2R16 amino acid sequence

MNGVLQVTFIVILSVEFIIGIFGNGFIAVVNIKDLVKGRKISSVDQILTALAISRIALLW
LILVSWWIFVLYPGQWMTDRRVSIMHSIWTTFNQSSLWFATSLSIFYFFKIANFSNPIFL
YLKVRLKKVMIGTLIMSLILFCLNIIIMNAPENILITEYNVSMSYSLILNNTQLSMLFPF
ANTMFGFIPFAVSLVTFVLLVFSLWKHQRKMQHSAHGCRDASTKAHIRALQTLIASLLLY
SIFFLSHVMKVWSALLLERTLLLLITQVARTAFPSVHSWVLILGNAKMRKASLYVFLWLR
CRHKE

SEQ ID NO:136

Mouse T2R16 nucleotide sequence

TTTATGATGGAAAGAATAAAACCATTAGCAAGGCTTAATGGCTTGTTTGGTATTAGACCT
GTACATTGTTTATGGAACATGATATGGAGCTTTGTTTATTGAATATGCACAATATTTTAG
AAGCATGTTTCAAAGAATCTTAAGTAATTACAATAGAAATTGAAGCATCCAAGTGAAGAT
GAATGGTGTCCTACAGGTTACATTTATAGTCATTTTGAGTGTGGAATTTATAATTGGCAT
CTTTGGCAATGGATTCATAGCGGTGGTGAACATAAAGGACTTGGTCAAGGGAAGGAAGAT
CTCTTCAGTGGATCAGATCCTCACTGCTCTGGCCATCTCCAGAATTGCACTGCTGTGGTT

AATATTAGTAAGT●GGTGGATATTTGTGCTTTACCCAGGACAAT●ATGACTGATAGAAG
AGTTAGCATAATGCACAGTATATGGACAACATTCAACCAGAGTAGTCTCTGGTTTGCTAC
AAGTCTCAGCATCTTTTATTTTTTCAAGATAGCAAATTTTTCCAACCCTATTTTTCTTTA
TTTAAAGGTCAGACTTAAAAAAGTCATGATAGGGACATTGATAATGTCTTTGATTCTCTT
TTGTTTAAATATTATCATTATGAATGCACCTGAGAACATTTTAATCACTGAATATAATGT
ATCTATGTCTTACAGCTTGATTTTGAATAACACACAGCTTTCTATGCTGTTTCCATTTGC
CAACACCATGTTTGGGTTCATACCTTTTGCTGTGTCACTGGTCACTTTTGTCCTTCTTGT
TTTCTCCCTGTGGAAACATCAGAGAAAGATGCAACACAGTGCCCATGGATGCAGAGATGC
CAGCACTAAGGCCCACATCAGAGCCTTGCAGACATTGATTGCCTCCCTCCTCCTGTATTC
CATTTTCTTCCTGTCTCATGTTATGAAGGTTTGGAGTGCTCTGCTTCTGGAGAGGACACT
CCTGCTTTTGATCACACAGGTTGCAAGAACAGCTTTTCCGTCAGTGCACTCCTGGGTCCT
GATTCTGGGCAATGCTAAGATGAGAAAGGCTTCTCTCTATGTATTCCTGTGGCTGAGGTG
CAGGCACAAAGAATGAAACCCTACAGTGTACAGACCTGGGGTATATTTATGTGGATGATC
TTACATATCTTAGAGGAAAATGGATTAAAAGAAATTCTCATATTTATAAATTTTTAGGTC
TGAATTACATAAAAATGTATATAATATTTTCAAAGTACAAGATAGTAGTTTATAACTTAC
ATGATAAATACTGTCTATGCATCTTCTAGTCTTTGTAGAATATGTAAAAACATGTT

SEQ ID NO:137

Mouse T2R17 amino acid sequence

MKHFWKILSVISQSTLSVILIVELVIGIIGNGFMVLVHCMDWVKKKKMSLVNQILTALSI
SRIFQLCLLFISLVINFSYTDLTTSSRMIQVMYNAWILANHFSIWIATCLTVLYFLKIAN
FSNSFFLYLKWRVEKVVSVTLLVSLLLLILNILLTNLETDMWTNEYQRNISCSFSSHYYA
KCHRQVLRLHIIFLSVPVVLSLSTFLLLIFSLWTHHKRMQQHVQGGRDARTTAHFKALQT
VIAFFLLYSIFILSVLIQIWKYELLKKNLFVVFCEVVYIAFPTFHSYILIVGDMKLRQAC
LPLCIIAAEIQTTLCRNFRSLKYFRLCCIF

SEQ ID NO:138

Mouse T2R17 nucleotide sequence

GAATTCTGGTCTGGCACCCCTGAGCTGTGTGAGTAGACACATTATCATGGAAAGAGATTC
AGAATCTGTCACTGTCAAAACTGCATGTTTGCTCCTCTGTTAGTGTGTTGGGGAAAGTTA

AGAAAAATACATT●ATGAGAATCAACTCAGAGGTTGTCAGAAA●GTCGAAACAGCATT
TTAAAAATTTACATCTCAACTGGATATATGAGCAAGTCTTTATAACTGATATATAAAATG
AAGCACTTTTGGAAGATATTATCTGTTATCTCCCAGAGCACACTTTCAGTCATTTTAATC
GTGGAATTAGTAATTGGAATTATAGGAAATGGGTTCATGGTCCTGGTCCACTGTATGGAC
TGGGTTAAGAAAAAGAAAATGTCCCTAGTTAATCAAATTCTTACTGCTTTGTCAATCTCC
AGAATTTTTCAGCTCTGTTTATTGTTTATAAGTTTAGTAATCAACTTTTCATATACAGAT
TTAACTACAAGTTCAAGGATGATACAAGTCATGTACAATGCTTGGATTTTAGCCAACCAT
TTCAGCATCTGGATTGCTACATGCCTCACTGTCCTTTATTTTCTAAAGATAGCCAATTTT
TCTAACTCTTTTTTTCTTTATCTAAAGTGGAGAGTTGAAAAAGTAGTTTCAGTTACACTG
TTGGTGTCATTGCTCCTCCTGATTTTAAATATTTTACTAACTAACTTGGAAACCGACATG
TGGACAAATGAATATCAAAGAAACATATCATGCAGCTTCAGTTCTCATTACTATGCAAAG
TGTCACAGGCAGGTGTTAAGGCTTCACATTATTTTCCTGTCTGTCCCCGTTGTTTTGTCC
CTGTCAACTTTTCTCCTGCTCATCTTCTCCCTGTGGACACATCACAAGAGGATGCAGCAG
CATGTTCAGGGAGGCAGAGATGCCAGAACCACGGCCCACTTCAAAGCCCTACAAACTGTG
ATTGCATTTTTCCTACTATATTCCATTTTTATTCTGTCTGTCTTAATACAAATTTGGAAA
TATGAATTACTGAAGAAAAATCTTTTCGTTGTATTTTGTGAGGTTGTATATATAGCTTTT
CCGACATTCCATTCATATATTCTGATTGTAGGAGACATGAAGCTGAGACAGGCCTGCCTG
CCTCTCTGTATTATCGCAGCTGAAATTCAGACTACACTATGTAGAAATTTTAGATCACTA
AAGTACTTTAGATTATGTTGTATATTCTAGACAAAAATTAACTGATACAAATGTCTTTTG
TATTTTTCATTTTAAATATCCTTTAATTTTGACTGCATGAAATTGATTTCTGCTTGCAAT
TATCACTGATTAAAACTATTAATAATTTAACTAG

SEQ ID NO:139

Mouse T2R18 amino acid sequence

MVPTQVTIFSIIMYVLESLVIIVQSCTTVAVLFREWMHFQRLSPVETILISLGISHFCLQ
WTSMLYNFGTYSRPVLLFWKVSVVWEFMNILTFWLTSWLAVLYCVKVSSFTHPIFLWLRM
KILKLVLWLILGALIASCLSIIPSVVKYHIQMELVTLDNLPKNNSLILRLQQFEWYFSNP
LKMIGFGIPFFVFLASIILLTVSLVQHWVQMKHYSSSNSSLKAQFTVLKSLATFFTFFTS
YFLTIVISFIGTVFDKKSWFWVCEAVIYGLVCIHFTSLMMSNPALKKALKLQFWSPEPS

SEQ ID NO:140

Mouse T2R18 nucleotide sequence

GCGTGCTTCACAGAGCAGTATACTACAAAGCAAATGTCATTGCTGCCATTGTATATTTCT
CTAAAGACATTTCACATTTTATCTCCCTGTCCCATTGTGTGCAGAGCCCACACTTCAATC
AATCAATTCCTTAATTATAAGCTATTGTTTCATTATTTCATTTCCTACGTTTTTTTGCAT
TTTTACTAAAACTCCAAAGCAGACATTTTCTAATTATAATCCTACATGTAGTTAGAATTT
TAAAAATTATATACTATTTTCTTTGCACCACTGAGTTCAGTAGGTTTTGAAGGTTTATGC
TTAACAATTGAACATTTCATGTTAGATTATTCCTGCCTTCCTAATCTTGAATAATTAAAT
GTCCATCCAGGCTTAGAATTCACAGAGTCAACAGCTTTCACCTTGATTCTCTCACTATCT
ATCAATGACTAGAATCTGTCTGTCACTTTTGAAACCGCTAATTAAATAGTTGGTGCTTAT
TTAAAGGGTGCCCCATGCCAAGAGAAAATGTATTTCTTCTCTAGATGCCTTCGTCCTTTA
CAAGTTACATGCTTTACTGATGGTGAATTGGTTTTCTTCCAGTTCATCTGGGTTAAGTGA
CCTAAGAACCTAGCCATGGAAGGAGAAACAGAAGCAAATATTAACGATACAAGAACAAGT
TCCAGAACATTGGAAAGTACTTAGTAAAGGCATTGGAATTAGCAAAAGAATAGTAGCGAA
GCAAAAAATACTTCATCTCCATTGGGAGGTCAAGAAAGACTATGCAGTGTTTTTGATGCA
ACTTGTCATCTCTGAGTTAGACGATTCAGCACACACTTTTGAGATTGAACTTCAACAGGT
GGAGCCAGCAGACCTGAGCTTTAGGAATGATGGTGGAATTTCCAAGCAAAGACTTCCGTT
ACCTTTTTGATGTCCCCTAACAATTCGGTTGCAATGCTCACACCGCCCAACTGTTGAAAT
GCTTGGGAAAAGGGATTCTGAGACTGGCATTAGTATGTCATTTGACAGAATGGAAACATT
GCCCAGGGCATTAATGCACAGTAAAGGATTCACCTTTTCTAAGTGCTCAAATTTTAAATT
TGnATATTTTTAGAAGACATTATTTAAAAGAAAGGTGGAGAGGATATCCAAACAGCACCT
TGAGCAGATAAAGAGGTGAAGAAGAAAAAACAACATGCGTACATGATGGATTTCTCTTTA
TGAAAATGATCAAATGATCTTAGGATCAAGAATCCACACCTGAATGAGATTTGCTTGTAT
CCCTGTGTGAATTTGACCTAACAAGCAAAGCACAGACAAATGCTGTAGATAGGGAAATGT
CTATGTCAAATGTGTGTAAGGAGGATTTGCATCCACAAAGAAGTGCCCTCTTATACTGAG
AGTGCTAAGAACACATGTCCGTTTCATATTCGGAAAGTGGTATAGAGCTGTTGAGTCTTT
GGCTAGGAAGAGACTTCAGAGTGGAAGC**ATGGTGCCAACGCAAGTCACCATCTTCTCCAT**
CATCATGTATGTGCTTGAGTCCTTAGTAATAATTGTGCAAAGTTGCACAACGGTTGCAGT
GCTATTCAGAGAGTGGATGCACTTTCAAAGACTGTCACCGGTGGAGACGATTCTCATCAG
CCTGGGCATCTCACATTTCTGTCTACAGTGGACATCAATGCTATACAACTTTGGTACTTA
TTCTAGGCCTGTCCTTTTATTTTGGAAGGTATCAGTCGTCTGGGAGTTCATGAACATTTT
GACATTCTGGTTAACCAGTTGGCTTGCTGTCCTCTACTGTGTCAAGGTCTCTTCCTTCAC
TCACCCCATCTTCCTCTGGCTGAGGATGAAAATCTTGAAACTGGTTCTCTGGTTGATACT
GGGTGCTCTGATAGCTTCTTGTTTGTCAATCATCCCTTCTGTTGTTAAATATCACATCCA

GATGGAATTAGT[redacted]CCCTAGATAATTTACCCAAGAACAATTCTT[redacted]ATTCTAAGACTACA
ACAGTTTGAATGGTATTTTTCTAATCCTTTAAAAATGATTGGCTTTGGTATTCCTTTCTT
CGTGTTCCTGGCTTCTATCATCTTACTCACAGTCTCATTGGTCCAACACTGGGTGCAGAT
GAAACACTACAGCAGCAGCAACTCCAGCCTGAAAGCTCAGTTCACTGTTCTGAAGTCTCT
TGCTACCTTCTTCACCTTCTTCACATCCTATTTTCTGACTATAGTCATCTCCTTTATTGG
CACTGTGTTTGATAAGAAATCTTGGTTCTGGGTCTGCGAAGCTGTCATCTATGGTTTAGT
CTGTATTCACTTCACTTCACTGATGATGAGCAACCCTGCATTGAAAAAGGCACTGAAGCT
GCAGTTCTGGAGCCCAGAGCCTTCCTGAGGCAGGAAACACAGTTAAGCCTCTAGGGTAAG
GAGACTTTGCATTGGCACAGTCCCTATAGTGTAATGCAAACTTGAACACAAACTTCATCC
CTTTTCACATCCACAAATGGCTGCATCTATACATCATCACCAGTCTTCCCTGTATTCTGA
CCCATTCTCTTCCTGTCCTATCCATAGTCCCCAGGTTGGTTTTGATTTTTCTCATGATCA
CACCAACTCTGCTTAGCTTTTGCCACCACTGTAATAGTAAACATGGGGTGTTCTATATAT
TACAGTCAAAATCATTCTCACATTGTTGATTGCCTCACAAATTCATATAAATCCCCCTTC
CTGTCAGGAATTTATTGTCTGCTCACTTAATGCTCACCATATATTAAAGCCATTAATTCC
CCCTTCCTACCTTGAGTTTAAGAAGGAAAATGTCTTACCATTGCCCACAACCTATTCTGC
TGCTTCTAGACTTTTATGCAAGTGATTTATACACACACACACACACACACACACATAC
AAACAAC

SEQ ID NO:141

Mouse T2R19 amino acid sequence

MMEGHMLFFLLVVVVQFLTGVLANGLIVVVNAIDLIMWKKMAPLDLLLFCLATSRIILQL
CILFAQLGLSCLVRHTLFADNVTFVYIINELSLWFATWLGVFYCAKIATIPHPLFLWLKM
RISRLVPWLILASVVYVTVTTFIHSRETSELPKQIFISFFSKNTTRVRPAHATLLSVFVF
GLTLPFLIFTVAVLLLLSSLWNHSRQMRTMVGTREPSRHALVSAMLSILSFLILYLSHDM
VAVLICTQGLHFGSRTFAFCLLVIGMYPSLHSIVLILGNPKLKRNAKTFIVHCKCCHCAR
AWVTSRNPRLSDLPVPATHHSANKTSCSEACIMPS

SEQ ID NO:142

Mouse T2R19 nucleotide sequence

CTGCAGCCTAGAC●CTAATGCATAGGAAACTTATATTCCCACC●CGTGACGTCACTCT
GACAGAAGTGAACTTATATTCCCACCTCCGTGACGTCACTCTGACAGAAGTGACTTGTTT
TTGTATGATGCTCCAGGATGCCTCATTAGCATTGAGGACAATCATAATTAAGTAAGGCAA
GGCATGAAGGTGGTCCTCACTAGGTACCTGGAGGCTTCTGGTTGCATGATTTACTTGTGA
TGACTCTGACACTTAAGAAGACCTGAAAAATGCAAAAGCTGTCATAAGGCACAGTTCGTT
TCTATGGTATCTCTTCCTTATTTGACTGACATTGAGTTGAGAAGGCAGCACTATAAACAA
ATGGGCCCCACCTTCCTCTTCCATTGTCTTTGGGTTGGCATCATCTCCAAAGGAACCTTG
GTCTAGTTGAAAGAAGCCAGAAATCATACATGGCTGAGACTGTGCATAACTCTATGTATC
ATTTAAAGAAGTCATTGGTTCTTCTTATTTTAAA**ATGATGGAAGGTCATATGCTCTTCTT**
CCTTCTGGTCGTGGTAGTGCAGTTTTTAACTGGGGTCTTGGCAAATGGCCTCATTGTGGT
TGTCAATGCCATCGACTTGATCATGTGGAAGAAAATGGCCCCACTGGATCTGCTTCTTTT
TTGCCTGGCGACTTCTCGGATCATTCTTCAATTGTGTATATTGTTTGCACAGCTGGGTCT
ATCCTGTTTGGTGAGACACACGTTATTTGCTGACAATGTTACCTTTGTCTACATTATAAA
CGAACTGAGTCTCTGGTTTGCCACATGGCTTGGTGTTTTCTACTGTGCCAAGATTGCTAC
CATCCCTCACCCACTCTTTCTGTGGCTGAAGATGAGGATATCCAGGTTGGTGCCATGGCT
GATCCTGGCATCTGTGGTCTATGTAACTGTTACTACTTTCATCCATAGCAGAGAGACTTC
AGAACTTCCTAAGCAAATCTTTATAAGCTTTTTTTCTAAAAATACAACTCGGGTCAGACC
AGCGCATGCCACACTACTCTCAGTCTTTGTCTTTGGGCTCACACTACCATTTCTCATCTT
CACTGTTGCTGTTCTGCTCTTGTTGTCCTCCCTGTGGAACCACAGCCGGCAGATGAGGAC
TATGGTGGGAACTAGGGAACCTAGCAGACATGCCCTCGTCAGTGCGATGCTCTCCATTCT
GTCATTCCTCATCCTCTATCTCTCCCATGACATGGTAGCTGTTCTGATCTGTACCCAAGG
CCTCCACTTTGGAAGCAGAACCTTTGCATTCTGCTTATTGGTTATTGGTATGTACCCCTC
CTTACACTCGATTGTCTTAATTTTAGGAAACCCTAAGCTGAAACGAAATGCAAAAACGTT
CATTGTCCATTGTAAGTGTTGTCATTGTGCAAGAGCTTGGGTCACCTCAAGGAACCCAAG
ACTCAGCGACTTGCCAGTGCCTGCTACTCATCACTCAGCCAACAAGACATCCTGCTCAGA
AGCCTGTATAATGCCATCTTAATTGTCCAACCTGAGGCTTAATCATTTCAAAGGGTAAAT
TGATGATCAAAGCCCAACACATGATATGACATCAAGGTCCATATCCCAGTAGTCATGTGG
AAATACCACCTTGCAAAATGATGTCATTGAGAAACCAGGGCAAATGGAGTCTAGGTCTTT
CAGTATGATTTGCTGCAG

SEQ ID NO:143

Mouse T2R20 amino acid sequence

MNLVEWIVTIIM[redacted]EFLLGNCANVFITIVNFIDCVKRRKISSA[redacted]IITAIAIFRIGLLWA
MLTNWHSHVFTPDTDNLQMRVFGGITWAITNHFTTWLGTILSMFYLFKIANFSNSLFLHL
KRKLDNVLLVIFLGSSLFLVAYLGMVNIKKIAWMSIHEGNVTTKSKLKHVTSITNMLLFS
LINIVPFGISLNCVLLLIYSLSKHLKNMKFYGKGCQDQSTMVHIKALQTVVSFLLLYATY
SSCVIISGWSLQNAPVFLFCVTIGSFYPAGHSCILIWGNQKLKQVFLLLLRQMRC

SEQ ID NO:144

Mouse T2R20 nucleotide sequence

CTAGATGGGCTGTTTCATATAATGACTGGAACTCCCTACATGCTCCACGTCTTGAGTTCT
AAAATTTCACTAACAAATTTTTGACTGCCATAAATAATGAAGGTTTAAAGAAAGAACAAC
ATTTGAAGCAATGGACCAGAATTCCTCTTTATTTGACTCTTAGCAAATTGGAATGCAGCA
TCCTTTCAAGAGCAGCACTGAAATATACCAGTCAATGGCAGAGAGTAAAAAAGTATGCAA
TTGGAGACATTATGGTAATATAAATTTCCATTAAAAATGAGACTGCATTCACCTATTACA
ACACATTGCTATTCTGCTCAACACAGAGTTAAAAAGAAACAAGAACTCTTGTATACATTC
AGTTAGTCACAAGTATAATTATGTTCACATATTTTAAAAAAATGAATCATGATCTGTGAA
TTGAGCCTGGCTTTTTTTGTCTCTCTCTTTTTATTCTTTTCCTTTAGACAGACACA**ATGA**
ATTTGGTAGAATGGATTGTTACCATCATAATGATGACAGAATTTCTCTTAGGAAACTGTG
CCAATGTCTTCATAACCATAGTGAACTTCATCGACTGTGTGAAGAGAAGAAAGATCTCCT
CAGCTGATCGAATTATAACTGCTATTGCCATCTTCAGAATTGGTTTGTTGTGGGCAATGT
TAACGAACTGGCATTCACATGTGTTTACTCCAGACACAGACAATTTACAAATGAGAGTTT
TCGGTGGAATTACCTGGGCTATAACCAACCATTTTACCACTTGGCTGGGGACCATACTGA
GCATGTTTTATTTATTCAAGATAGCCAATTTTTCCAACAGTCTATTTCTTCATCTAAAAA
GAAAACTTGACAATGTTCTACTTGTGATTTTCCTGGGATCGTCTCTGTTTTTGGTTGCAT
ATCTTGGGATGGTGAACATCAAGAAGATTGCTTGGATGAGTATTCATGAAGGAAATGTGA
CCACAAAGAGCAAACTGAAGCATGTAACAAGCATCACAAATATGCTTCTCTTCAGCCTGA
TAAACATTGTACCATTTGGTATATCACTGAACTGTGTTCTGCTCTTAATCTATTCCCTGA
GTAAACATCTCAAGAATATGAAATTCTATGGCAAAGGATGTCAAGATCAGAGCACCATGG
TCCACATAAAGGCCTTGCAAACTGTGGTCTCTTTTCTCTTGTTATATGCCACATACTCTT
CCTGTGTCATTATATCAGGTTGGAGTTTGCAAAATGCACCAGTCTTCCTGTTTTGTGTGA
CAATTGGATCCTTCTACCCAGCAGGTCATTCTTGTATCTTGATTTGGGGAAACCAGAAAC
TTAAACAGGTCTTTCTGTTGTTGCTGAGGCAGATGAGATGCTGACTGAAAAAATGAAAGT
CCCCCTGTCTCTAG

SEQ ID NO:145

Mouse T2R21 amino acid sequence

MGSNVYGILTMVMIAEFVFGNMSNGFIVLINCIDWVRKGTLSSIGWILLFLAISRMVLIW
EMLITWIKYMKYSFSFVTGTELRGIMFTWVISNHFSLWLATILSIFYLLKIASFSKPVFL
YLKWREKKVLLIVLLGNLIFLMLNILQINKHIEHWMYQYERNITWSSRVSDFAGFSNLVL
LEMIVFSVTPFTVALVSFILLIFSLWKHLQKMHLNSRGERDPSTKAHVNALRIMVSFLLL
YATYFISFFLSLIPMAHKTRLGLMFSITVGLFYPSSHSFILILGHSNLRQASLWVMTYLK
CGQKH

SEQ ID NO:146

Mouse T2R21 nucleotide sequence

CTCTTTTGAAGACAATAGTTGTTCTACTAGCTATTGATAGCATGTTTACATTTGTCATTT
TCAAGTATGTTCAGAAACAAAGCTACATATTGTGGGGAGTATATAAAATATGAAAGCATG
CCATTCCCAGGCATCCAAGGATCCCTGTGTATTAAAAGGCAACAAAGCAGAACCAAATGT
TCTGTTTTGGACATGAGCTTCTTCCAATTCAACTGCTGAAAAATTTGGATAACTACATAT
AAAACTAAGAACACAGAGTGTCACAGAGCAGTCTCTGCTCTCCAATTCACCAGGATTAAT
ATTGACAGACCCAAAAGATGTCATTTAGGTAAATTTTGGATGAATCATATTGTTGTCACC
TTTGTGCTCTAGAACATAAGCTGATAGAATCAAATTTTCTTTAGCAGAGACAATGCAAAT
TGATATAACAGTGAAAGAGAATATATCTTTATTTGCATGTTAGCAAATGACAGCTGGATG
CACTTCATGATTTTCTGCAATCTAGTTCAGTCTTTAGAAGGATATATATATATATATATA
TATATATATATATATATATATATATATATATATATAAACCTTAGTCTTGAAAGATATCAGAA
AGAAGGATTTCACAAGAATGTACAGAGCCATTAGCAAAATTTTAATATACTCATCGACAT
TAGGTCAGTCACTACATAAGAAGGACTTGAATGAAAGCTTATCTTAGTTTTTGAGACTAC
AGGGACATTTCACCTTGCCAAATGAGAAGCAGTGAGTCTTCTTTGTCTGGAC**ATGGGAAG**
CAATGTGTATGGTATCTTAACTATGGTTATGATTGCAGAGTTTGTATTTGGAAATATGAG
CAATGGATTCATAGTGCTGATAAACTGCATTGATTGGGTCAGGAAAGGAACTCTTTCTTC
CATTGGTTGGATCCTGCTTTTCTTGGCCATTTCAAGAATGGTGTTGATATGGGAAATGTT
AATAACATGGATAAAATATATGAAGTATTCATTTTCATTTGTGACTGGAACAGAATTACG
GGGTATCATGTTTACCTGGGTAATTTCCAATCACTTCAGTCTCTGGCTTGCCACTATTCT

CAGCATCTTTTAT[redacted]TGCTCAAAATAGCCAGTTTCTCCAAACCGG[redacted]TTTCTCTATTTGAA
GTGGAGAGAGAAGAAAGTGCTTCTGATTGTCCTTCTGGGAAATTTGATCTTCTTGATGCT
CAACATATTACAAATAAACAAACATATAGAACACTGGATGTATCAATATGAGAGAAATAT
AACTTGGAGTTCTAGAGTGAGTGACTTTGCAGGGTTTTCAAATCTGGTCTTATTGGAGAT
GATTGTGTTCTCTGTAACACCATTCACAGTGGCCCTGGTCTCCTTCATCCTGTTAATCTT
CTCCTTGTGGAAACATCTACAGAAAATGCATCTCAATTCTAGAGGGGAACGAGACCCCAG
CACTAAAGCCCATGTGAATGCCTTGAGAATTATGGTCTCCTTCCTCTTACTCTATGCCAC
TTACTTCATATCTTTTTTTCTATCATTGATTCCCATGGCACATAAAACACGACTGGGTCT
TATGTTTAGCATAACTGTTGGGCTTTTCTACCCTTCAAGCCACTCATTTATCTTAATTTT
GGGACATTCTAATTTAAGGCAAGCCAGTCTTTGGGTGATGACATATCTTAAATGTGGGCA
AAAGCATTAGAATTTCACTATTCCATAAGGCAGCCAAACCACGTGCTACTAGGTATATGA
TACTACTCAGTGGTAAAGCCCTAGGCAAACATTAACCTTAGAAAATATATAATTTTGTGA
CTCTTCTGTATTTGATAAATCACTCACATATTTAGAAGAATGCTACAGTAGTGTGATCTT
GTACATGATTGTAACAATTCAATTTTATTAATATAGTTCAGGCATGATAACATACCCCTG
ATAACTGAAAAGTAAGTAGGATGCTACATATATATTTAGATCTAGACTTAGGGGCAAAGA
GAGACCCAGCTGATAGCTGTGCAATAAAGATTTTAATTTTCATCCTGTTGTGAGTTATCT
GAAATCTATGTCACTGAAGGCATAAGCAAGATTTTCACACACTGAAACAATCTCTTATGC
TTTCTTATATTGTTTTAAAAGTAAATTAGAAAATTTAAATAAACTTAATGGCAATTGAAA
TTACAAAAGCTAAACACATGTGGTTATTAGAAATTAGACTGTATGTAGGTCCTAGGGGAT
GGCTTAGTAAAGTGCTTTGTTGCAAGCTTCAGGATATGATTCTAAATCCCTAGATTCAAT
TAAAAACCTGGCATAAATAGCCAATGTAAAATTTGTCTGTAAAATGTAACCAGTGCTAAG
AGTACCAAGACAACAAAATGTTTACTTTTAAAACCATTTATTGATATTCTTTTAAAAATA
GGTATGTATTTTACTATTTAAATAAGATTTTGTCAAAAGCTAGTCTTGACACCTTAGGTA
AACATAGGAAGGCAACAAGTTTGAAGTCAGCTACTGGGGACAGTGCTGCTAGCAGCTGAC
AGAGGCCACTGCTGACTACAGCAGATCATTTACAGGTTCAGCACTAG

SEQ ID NO:147

Mouse T2R22 amino acid sequence

MSSLLEIFFVIISVVEFIIGTLGNGFIVLINSTSWFKNQKISVIDFILTWLAISRMCVLW
TTIAGASLRKFYKTLSYSKNFKFCFDIIWTGSNYLCIACTTCISVFYLFKIANFSNSIFF
WIKQRIHAVLLAIVLGTLMYFILFLIFMKMIANNFIYKWTKLEQNTTFPVLDTLSGFLVY
HSLYNGILIFFFIVSLTSFLLLIFSLWSHLRRMKLQGIHTKDISTEAHIKAMKTMMSFLL

FFIIYYISNIML[redacted]ASSILDNVVAQIFSYNLIFLYLSVHPFLL[redacted]NSKLKWTFQHVLRK
LVCHCGGYS

SEQ ID NO:148

Mouse T2R22 nucleotide sequence

AAATGAATAATTTCATGCAAAGGATACCATTAGAATATGATCACTATTTAAATTTTAGCA
AATACATATTCAAATACCAGCACAATGTTTCAAATTTAAAATATAAACATTATAAAACCC
AGCAGAGAACAAAATGATAGCCTTGATAATTGTTGGTTTGCTCAAGAAAAATGGGTGTAT
ACTTTAACATTTAATTGGGAACTCAGTTGAGAGCATACATTTAGGGTTTTACAGAGGTAT
TCATTGCCCATTTAAGATTTGGATTCACACATCTACATCAATGTGGCTGTAATCCATTTT
CCCATGATGAAATAAGGTAGAGACTGCCTATTAAACGAC**ATGTCGAGCCTACTGGAGATT**
TTCTTTGTGATCATTTCGGTTGTAGAATTCATAATAGGAACTTTGGGAAATGGATTTATT
GTCCTGATAAACAGTACTTCTTGGTTCAAGAATCAGAAAATCTCTGTAATTGATTTCATT
CTTACTTGGTTGGCCATCTCCAGAATGTGTGTTCTATGGACAACAATTGCTGGTGCCTCT
CTCAGGAAATTCTACAAGACGTTAAGTTACTCTAAGAATTTCAAATTTTGTTTTGACATT
ATCTGGACAGGATCCAACTATTTATGCATAGCCTGTACAACGTGCATCAGTGTCTTCTAC
TTGTTCAAGATTGCCAACTTTTCTAATTCCATTTTCTTCTGGATTAAACAGAGAATTCAT
GCAGTACTTCTGGCTATTGTCCTAGGCACACTCATGTATTTCATTTTATTTCTCATTTTT
ATGAAAATGATAGCTAATAATTTTATCTACAAATGGACAAAATTGGAACAAAACACAACA
TTCCCTGTTTTAGATACTCTAAGTGGTTTCTTAGTCTACCATAGCCTCTACAATGGGATT
CTCATTTTCTTTTTTATAGTGTCTCTGACCTCATTTCTTCTTTTAATCTTCTCTTTATGG
AGCCACCTTAGGAGGATGAAACTACAGGGCATACATACCAAAGACATAAGCACAGAAGCA
CACATAAAAGCTATGAAAACTATGATGTCATTCCTTTTGTTCTTCATCATATATTATATT
AGCAACATTATGCTTATTGTGGCAAGCTCCATTCTTGACAATGTGGTTGCACAAATTTTC
TCTTATAACCTAATATTTCTGTATTTATCTGTTCATCCTTTTCTTCTGGTTTTATGGAAC
AGCAAATTGAAATGGACATTCCAGCATGTATTGAGAAAGCTGGTGTGTCATTGTGGAGGT
TATTCTTGATTTCAGTAAATACACTCAATATAACTGATGGATTTCTAAGGTAAGAAAAAT
GGAACAAGGAATAAAGAGGAGAAATATATTCCTTTTCAGATCATCTGCTCTGTCATTCTG
TCCTTAGCATGCTATTAAGAATTGTTGACTAAATCCAGTCATTTTTAACATGAGGAAAGG
ATGTTTCAATCCAACTTAGAGAGGGTACAAAATAGTCCTAGGAGGCAG

SEQ ID NO:149

Mouse T2R23 amino acid sequence

MFSQKINYSHLFTFSITLYVEIVTGILGHGFIALVNIMDWVKRRRISSVDQILTALALTR
FIYVLSMLICILLFMLCPHLPRRSEMLSAMGIFWVVNSHFSIWLTTCLGVFYFLKIANFS
NSFFLYLKWRVKKVILIIILASLIFLTLHILSLGIYDQFSIAAYVGNMSYSLTDLTQFSS
TFLFSNSSNVFLITNSSHVFLPINSLFMLIPFTVSLVAFLMLIFSLWKHHKKMQVNAKQP
RDVSTMAHIKALQTVFSFLLLYAIYLLFLIIGILNLGLMEKIVILIFDHISGAVFPISHS
FVLILGNSKLRQASLSVLPCLRCQSKDMDTMGL

SEQ ID NO:150

Mouse T2R23 nucleotide sequence

AATTTTCAGCAACCAATATGTAGACTGCTTAAATGCATCAGAAACATTATAAATTGAAGC
ATGTTTTCACAGAAAATAAACTACAGCCATTTGTTTACTTTTTCAATCACCTTGTATGTG
GAAATAGTAACGGGAATCTTAGGACATGGATTCATAGCATTAGTGAACATCATGGACTGG
GTCAAAAGAAGAAGGATCTCTTCAGTGGATCAGATTCTCACTGCTTTGGCCCTTACCAGA
TTCATTTATGTCTTGTCTATGCTGATTTGCATATTGTTATTCATGCTGTGCCCACATTTG
CCTAGGAGATCAGAAATGCTTTCAGCAATGGGTATTTTCTGGGTAGTCAACAGCCATTTT
AGCATCTGGCTTACTACATGCCTCGGTGTCTTTTATTTTCTCAAGATAGCCAATTTTTCT
AACTCTTTTTTTCTTTATCTAAAGTGGAGAGTTAAAAAAGTGATTTTAATAATAATCCTG
GCATCACTGATTTTCTTGACTTTACACATTTTATCTTTAGGGATATATGATCAGTTCTCA
ATTGCTGCTTATGTAGGAAATATGTCTTATAGTTTGACAGATTTAACACAATTTTCCAGT
ACTTTCTTATTCTCCAACTCATCCAATGTTTTCTTAATCACCAACTCATCCCATGTTTTC
TTACCCATCAACTCCCTGTTCATGCTCATACCCTTCACAGTGTCCCTGGTAGCCTTTCTC
ATGCTCATCTTCTCACTGTGGAAGCATCACAAAAAGATGCAGGTCAATGCCAAACAACCT
AGAGATGTCAGTACTATGGCCCACATTAAAGCCTTGCAAACTGTGTTCTCCTTCCTGCTG
CTGTATGCCATATACTTACTTTTCCTTATCATAGGAATTTTGAACCTTGGATTGATGGAG
AAAATAGTGATACTGATATTTGACCACATTTCTGGAGCAGTTTTTCCTATAAGCCACTCA
TTTGTACTGATTCTGGGAAACAGTAAGCTGAGACAAGCCAGTCTTTCTGTGTTGCCTTGT
CTAAGGTGCCAGTCCAAAGATATGGACACCATGGGTCTCTAGTAAATTCCAGAGTACATT
TTGTAAAAATCTTGAGGATGATCAGTTCATAGAAAAAAGTTACCTTATGGGGGAAAATAA
AAAGTGGGGCTTCAATCCTGGGAGTAATAATACACAGGAGGGTAGGACAGCATGAAGGAG

ACTAGCACTATAT[redacted]AGTGGTCTCATACAGGATATGGGAAAGGAA[redacted]ATTTATGCAATAAA
GAGGGAGATCATATTGGAGGATGAGGAGGCATTACATATGTAAAATGACTATAAGAATGG
AATCATGCTAATCTAAAAAAATCTGTAATGCATTTCATTCAGACTATATACATATATGCC
TATATATGGATATATGGGGATATATATTCTATACATATTTTAAAAGAACCTTTCTTATAT
AG

SEQ ID NO:151

Mouse T2R24 amino acid sequence

MVPVLHSLSTIILIAEFVWGNLSNGLIVLKNCIDWINKKELSTVDQILIVLAISRISLIW
ETLIIWVKDQLISSITIEELKIIVFSFILSSHFSLWLATALSIFYLFRIPNCYWQIFLYL
KWRIKQLIVHMLLGSLVFLVANMIQITITLEERFYQYGGNTSVNSMETEFSILIELMLFN
MTMFSIIPFSLALISFLLLIFSLWKHLQKMPLNSRGDRDPSATAHRNALRILVSFLLLYT
IYFLSLLISWVAQKNQSELVHIICMITSLVYPSFHSYILILGNYKLKQTSLWVMRQLGCR
MKRQNTPTT

SEQ ID NO:152

Mouse T2R24 nucleotide sequence

CAAAGAGGAGAAATATTTAGCTACACAGTGTACCACATACAAGCCGTTCAATCAGTATAA
GGGGAGCAGTCATATAGAATTTGGGCTTTCTTTCTTTTAAT**ATGGTACCTGTTCTGCACA**
GTCTCTCCACCATCATACTAATTGCAGAGTTTGTTTGGGGAAATTTGAGCAATGGTTTGA
TAGTGTTGAAGAACTGCATTGACTGGATCAATAAAAAAGAGCTCTCCACAGTTGATCAAA
TACTCATTGTCTTGGCAATTTCAAGAATTAGTCTCATCTGGGAAACACTAATTATATGGG
TTAAAGATCAACTAATTTCATCTATTACTATTGAAGAATTAAAAATAATTGTGTTCAGCT
TTATACTATCTAGCCACTTCAGTCTCTGGCTTGCTACAGCTCTCAGCATCTTCTATTTAT
TCAGAATACCTAATTGCTACTGGCAGATCTTTCTCTACTTGAAATGGAGAATAAAGCAAC
TGATTGTCCACATGCTTCTGGGAAGCTTGGTGTTCTTGGTTGCAAATATGATACAGATAA
CCATCACTCTTGAAGAGAGGTTCTATCAATATGGAGGAAATACAAGTGTAAATTCCATGG
AGACTGAGTTCTCAATTTTGATAGAGCTGATGTTATTTAACATGACTATGTTCTCCATTA
TACCATTTTCATTGGCCTTAATTTCTTTTCTTCTGCTAATCTTCTCTTTATGGAAACATC
TCCAGAAGATGCCACTCAATTCTAGAGGAGATAGAGACCCTAGTGCTACGGCCCACAGAA

ATGCCTTGAGAAT●TGGTCTCCTTCCTCTTGCTCTATACTATA●TTTCCTGTCTCTTC
TTATATCATGGGTTGCTCAGAAGAATCAAAGTGAACTGGTTCACATTATTTGTATGATAA
CTTCACTCGTGTATCCTTCATTCCACTCATATATCCTGATTCTGGGAAATTATAAATTAA
AGCAGACCTCTCTTTGGGTAATGAGGCAGCTGGGATGTAGGATGAAAAGACAGAATACAC
CAACTACATAAGGCAGCCAAACAGTCTATTGGGTTTTAGATAACAAATCTAAATCTATGA
GGAAGTAGTTCAATAACATTTTTCCCCTTGACATGGAGTAGCAGGGTTTTTTTTTATTAG
ATATTTTCTTTACTTACATTTCAAATGCTATCCCGAAAATTCCCTGTACCCTCTCCCTGT
CCTGTTCCCCTACCCACCCACTCCCACTTCTTGGCCCTGGCATTCCCCTGGAGTATCAGT
TTTTTATTAGTCAAACTATCTCACTGACTAAGGGTCATAAAACAAGTTATTTTAACACTA
ATTTCAATTAAATCAAAGGTAAAGTGTCAGCACATGCCTTTAATCACACAATTCCATCAA
ATTCAGCACTCAGGAGAGGGTGATCTCTGTGAATTCCAGCACACTGGCGGCCGTTACTAG
TGGATCCGAGCTCGGTACCAAGCTT

SEQ ID NO:153

Mouse T2R25 amino acid sequence

MMGIAIDILWAAIIIVQFIIGNIANGFIALVNIIDWVKRRKISLMDKIITALAISRIYLL
WSTFLITLTSSLDPDIKMAVKIIRISNNTWIIANHFSIWFATCLSIFYFLKIANFSNYIF
LYLRWRFKKVVSVTLLISLIFLLLNILLMNMHIDIWSDKSKRNLSFSVRSNNCTQFPRLV
LLINTMFTSIPFTVSLLAFLLLIFSLWRHLKTMQYYAKGSEDTTTAAHIKALHMVVAFLL
FYTVFFLSLAIQYWTSGSQENNNLFYATIVITFPSVHSCILILRNSQLRQASLLVLWWLL
CKSKDVRMLVP

SEQ ID NO:154

Mouse T2R25 nucleotide sequence

AAAACTATTCGAATTGAACACAGTAACCAATTCTTCAGCGGACTTACACAAATCAAGCTA
TTATCTTATGGATGATGGGTATTGCCATAGATATCTTATGGGCAGCTATTATCATTGTGC
AATTCATAATTGGGAATATTGCAAATGGATTCATAGCATTGGTGAACATCATAGACTGGG
TGAAGAGAAGAAAAATCTCTTTAATGGATAAGATCATTACTGCTTTGGCAATCTCTAGGA
TTTATCTGCTGTGGTCTACATTCTTAATTACACTAACATCTTCACTGGATCCAGATATTA
AAATGGCTGTGAAAATCATTAGAATAAGCAATAACACCTGGATTATTGCAAATCATTTCA

GCATTTGGTTTG[redacted]ACATGTCTCAGCATCTTTTATTTTCTCAA[redacted]AGCCAATTTTTCTA
ACTATATTTTTCTCTACTTAAGGTGGAGATTTAAGAAGGTGGTTTCAGTGACATTGCTAA
TCTCTCTTATCTTCCTGCTTTTAAATATTTTACTGATGAACATGCATATTGATATCTGGA
GTGATAAGTCCAAAAGAAACCTTTCTTTTAGTGTCAGATCAAATAATTGCACTCAGTTTC
CCAGACTTGTCCTTTTAATCAACACAATGTTCACATCAATCCCCTTCACTGTGTCCCTGT
TGGCTTTTCTGCTTCTCATCTTCTCCCTGTGGAGACACCTGAAAACCATGCAATACTATG
CTAAAGGCTCCGAAGACACCACCACAGCTGCACATATAAAGGCCTTGCACATGGTAGTGG
CCTTTCTCCTGTTCTACACAGTTTTCTTTTTGTCTCTTGCCATACAATATTGGACCTCTG
GGTCTCAAGAGAATAACAACCTGTTTTATGCCACAATTGTAATTACTTTCCCTTCAGTCC
ATTCATGTATCCTGATTCTGAGAAACAGCCAGCTGAGGCAGGCATCTCTGTTGGTGCTGT
GGTGGCTGCTGTGCAAGTCCAAAGATGTACGGATGTTGGTTCCCTGAAATACTCTGTCAA
TGCTCTTTAGTAGTGAAGAAGAAAATAGCTTAGTTAAGGAAATTCTTGTTCATTACCGAA
GTATACTTTCAAGTTTATGTATC

SEQ ID NO:155

Mouse T2R26 amino acid sequence

MLPTLSVFFMLTFVLLCFLGILANGFIVLMLSREWLLRGRLLPSDMILFSLGTSRFFQQC
VGLVNSFYYFLHLVEYSGSLARQLISLHWDFLNSATFWFCTWLSVLFCIKIANFSHPAFL
WLKWRFPALVPWFLLGSILVSVIVTLLFFWGNHTIYQAFLRRKFTGNTTFKEWNRRLEID
YFMPLKVVTMSIPCSLFLVSILLLISSLRRHSLRMQHNTHSLQDPNVQAHSRALKSLISF
LVLYAVSFVSMIIDATVFISSDNVWYWPWQIILYFCMSVHPFILITNNLRFRGTFRQLLL
LARGFWVA

SEQ ID NO:156

Mouse T2R26 nucleotide sequence

GAATTCTAGACAAGGAAAGACACACACTAAATGACTTTACTTGTGGGACCTAAAATAACC
AAAATAAGTCAAAATCACAGTGATGTTACTAGGGATCTAGGATAAGGGAATGAAGAGAAA
GATGTTGGTCATAGAGTACAAAAATTCAGCTAAGAACTCAGTCCTGGAGGCTGAATGTAT
AGCTGTGTGACAGACAGCAGCTAGCCATACCAGAGTATACACTTGCCTCTTGCTGAAAGA
GTAGATCTTATGTGTCCTTGTCACACATAAAAGTAATTGAAAAAGTAACTCTCTGAGATG

ACAGATACGTTA[redacted]ATGGTTTTACTTTTCAACCTGCTCCAGTAC[redacted]GTCCCTTTAATGTT
TGTGCTAGTAGATGGGGGACTCTCAAGTATCTTTGTGGTAGACAAATCTAAGGTGGCCTT
CATGAATACCAACCCAGACTTTTGTGACTTTGTGATCCCCCACTTTTGAAGTGGATAAGA
GCTGTGACTTGAGTCTAATCAAAGGAGTCCAACGTGTTGTTTATTCTGTAACAGTGCTTT
GTGTTTCTAGTTAATAACACAGGCAAAGAAGGCTAGGGTGACATTCCTAGGATTGTGTTA
TTTCTATCTTGCTCATGCCTCCCTCTGCTGGTCTAATGAAATAAGTCAGTGGCCATATTT
AAATATGACTACGTGGCAAATACTGATGATAGCCTGTGTGTTCCAACAAATATCCAGTAG
GAGACCTAGGCATTCAGTCCTGCAGCCACAAGGAAATAGGTTCTTTCACTGGAAAAAGAG
CAGTTTAGATGGTTATAAATTACTTAATCCATAGAAGCCATAGGGGCTTTATGTAGAGAT
TTGGGTAGAGAGGTAGACCTAGATATTGACTTAGGAGTGGCTATTCCTGAGTGGGGGTAG
ATATATGGCAGGGAAACTCAGATAAGAAAGACTTCTTTAGTGTCACGATTTTTCCTAGGT
ATCTCCTTGTGCCAGATATCTATGCGTCTATGTACCTACCTACCTACCTACCTACCTACC
TACCTACCTACCTACTGACACCTAATAGGAAGAGGCAAGTGGTCACAACCTGCAATGATG
GGATAAGAATGATGGAACTCAGTTACCAAGATTAAAATACCTTCCCCACTGATGTTATTG
CAAGCATGGCAGCATGTAGGCAAAATCAGAGAAGGCAAATCATGAGCAGCTGCTGCCCCA
TGGTACCCGAGCCCGGGAAATATTTGCATCATATCTGAGCCAAAAGCACACCTTTTATCT
ACTGCCTGAGCATTTTTCACATTGAAGTTCTGGCTCACATGCAGAATCCAACCATTTATC
TCCTGTCTCCAGAAGGGAGTGTCAGGGACTGTGGGTAGGGGCAGGGAGGAGGCCAGGAAC
CAAGGCAATCAGTGGTGACAGGAGGAGGGACTGAA**ATGCTACCAACATTATCAGTTTTCT**
TCATGTTGACCTTTGTTCTGCTCTGTTTCCTGGGGATCCTGGCCAACGGCTTCATTGTGC
TGATGCTGAGCAGGGAATGGCTACTGCGTGGTAGGCTGCTCCCCTCGGACATGATCCTCT
TCAGTTTGGGCACCTCCCGATTCTTCCAGCAGTGTGTGGGATTGGTCAACAGTTTCTATT
ACTTCCTCCATCTGGTTGAGTACTCCGGGAGCCTTGCCCGGCAGCTCATTAGTCTTCACT
GGGACTTCTTGAACTCAGCCACTTTCTGGTTTTGTACCTGGCTCAGCGTCCTGTTCTGTA
TCAAGATTGCTAACTTCTCCCATCCTGCCTTCCTGTGGTTGAAGTGGAGATTCCCAGCGT
TGGTGCCCTGGTTCTTGTTGGGCTCTATCTTGGTGTCCGTCATTGTAACTCTGCTGTTCT
TTTGGGGAAACCACACTATATATCAGGCATTCTTAAGGAGAAAGTTTACTGGGAACACAA
CCTTTAAGGAGTGGAACAGAAGGCTGGAAATAGACTATTTCATGCCTCTGAAAGTTGTCA
CCATGTCAATTCCTTGTTCTCTTTTTCTGGTCTCAATTTTGCTGTTGATCAGTTCTCTCA
GAAGGCATTCGCTAAGAATGCAGCACAATACCCACAGCTTGCAAGACCCCAACGTCCAGG
CTCACAGCAGAGCCCTGAAGTCACTCATCTCATTCCTGGTTCTTTATGCGGTGTCCTTTG
TGTCCATGATCATTGATGCTACAGTCTTCATCTCCTCAGATAATGTGTGGTATTGGCCCT
GGCAAATTATACTTTACTTTTGCATGTCTGTACATCCATTTATCCTCATCACCAATAATC
TCAGGTTCCGCGGCACCTTCAGGCAGCTACTCCTGTTGGCCAGGGGATTCTGGGTGGCCT

AGAAGGCTTGGTC●TTTATCTAGAGCCTTTGAAGAGACTCAGG●AGGGTAACTTCACT
TGGAAGTGAGCTCATCTACGTGGAAATGTCTTTGTAGGCAGGCATGGGGTCATACTGTGA
GGTTCCTCATTGGGAAAGAGGAGAAGAAAATACAGAGTGTCCTTCCTTACCTTAGGATAT
TATGAAAGTGGAAATTCCGAATCCTGGACCAGTATTGATCTAAGTGCAAAGTACAATATG
TCCTGTTCCTTTCATGTCTGTTTTCCTTTTGTTACTGATTCATTCTCTAGGGAATAGTCT
TGATCAACTGAATCATCTCATCTGGCTGGCCACTGGGGAGGTAAAAGAACTTTGTGTCAC
TGCTGCATTGGGATATACATGGGTGGGAAGCAAGTGTCCCTGAGGCAGAGTAGCACTCAG
TATGAGAACCTCAAAGAGCAGGTGGCTGTGCATGCAGGGGCTGGGGCAAGGAGTCCTGAT
CACTCTTCACTGTATGGGGATTATTTGTCTCTTGCCAAAATTTGGAGACTTTGGCTTTAG
TTTTGTGAAGATGACTGGAAAAATTCTTAATGCTACCCTGTATCATTTCTCAATAATATT
TTCCTTTTCCTGCCTTTAATTTTCTCCTATCTGCAGCGCCCCTTGCTTGTTATCCGTAAA
TAAATAAATAAATAAATAAATAAGCCCAATCCTCATTTTCCTGTCTTTGGGAACCCTTTT
ACTTCCCCAGGTATACGCTACAAAGCCACTTCTGCATTGAATAAACATTATCTTTCATTC
AGAAAAAGACTTAAGAATCTCACCTTTACAAAAAAAAAAAAAAAAGAATCTCACTTATTT
TATATTCAAATTCCATTTTTAAAAAGAAAAGCACAGCATTAATTTTTCTAAATACTGTTT
ATAAAAATAACTTGCTCTAAGAATTATACAAATGTTTTGAAAGGTAACTTTGGAAAAAAA
GTGTGATTAGACATGGATGTTTGTAAGACAGAACAAAGAGCTCTTGGAAGTCCATGGCAG
CTCATTGGTCTTGCCTTCAGTAGAGCCTGTCTGAATCCTGTAACCTCTTATGCCCTTTTG
TAGCTTTTCTGCAGATC

SEQ ID NO:157

Mouse T2R27 nucleotide sequence

GAATTCGCCCTTGCGGGATCCGGGAACGGATTCATAGCACTGGTAAACTTCATGGGCTGG
ATGAAGAATAGGAAGATTGCCTCCATTGATTTAATCCTCACAAGTCTGGCCATATCCAGA
ATTTGTCTATTGTGCGTAATACTATTAGATTGTTTTATATTGGTGCTATATCCAGATGTC
TATGCCACTGGTAAAGAAATGAGAATCATTGACTTCTTCTGGACACTAACCAATCACTTA
AGTATCTGGTTTGCAACCTGCCTCAGCATTTACTATTTCTTCAAGATAGGTAATTTCTTT
CACCCACTTTTCCTATGCCTCAAGTCTAGACGCCAAGGGC

SEQ ID NO:158

Mouse T2R28 amino acid sequence

GREWLRYGRLLPLDMILISLGASRFCLQLVGTVHNFYYSAQKVEYSGGLGRQFFHLHWHF
LNSATFWFCSWLSVLFCVKIAN

SEQ ID NO:159

Mouse T2R28 nucleotide sequence

GAATTCGCCCTTGCGGGATCCGGGAACGGGTTTATTGTGCTGGTGCTGGGCAGGGAGTGG
CTGCGATATGGCAGGTTGCTGCCCTTGGATATGATCCTCATTAGCTTGGGTGCCTCCCGC
TTCTGCCTGCAGTTGGTTGGGACGGTGCACAACTTCTACTACTCTGCCCAGAAGGTCGAG
TACTCTGGGGGTCTCGGCCGACAGTTCTTCCATCTACACTGGCACTTCCTGAACTCAGCC
ACCTTCTGGTTTTGCAGCTGGCTCAGTGTCCTGTTCTGTGTGAAGATTGCTAACATCACA
CACTCCACCTTCCTGTGTCTCAAGTCTAGACGCCAAGGGCG

SEQ ID NO:160

Mouse T2R29 amino acid sequence

MDGIVQNMFTFIVIVEIIIGWIGNGFIALVNCIHWYKRRKISALNQILTALAFSRIYLLL
TVFTVIAVSTLYTHVLVTRRVVKLINFHLLFSNHFSMWLAACLGLYYFLKIAHFPNSIFV
YLKMRINQVVSGTLLMSLGLLFLNTLLINSYIDTKIDDYREHLLYDFTSNNTASFYRVIL
VINNCIFTSIPFTLSQSTFLLLIFSLWRHYKKMQQHAQRCRDVLADAHIRVLQTMVTYVL
LCAIFFLSLSMQILRSELLKNILYVRFCEIVAAVFPSGHSCVLICRDTNLRGTFLSVLSW
LKQRFTSWIPNINCRSSCIF

SEQ ID NO:161

Mouse T2R29 nucleotide sequence

AGCTTGATATTTCCTATTTGTTACTGCACAGAGTTTTTTTTAAAAATTGAGTTTGTTATG
TGGATTCAATACTCAGATAGAGCTCTTTAATTTTTTTACAGTGACCTCATGAATCATAAC
TTGCCTTACAGACA**ATGGATGGAATCGTACAGAACATGTTTACATTCATTGTAATTGTGG**
AAATAATAATAGGATGGATTGGAAATGGATTCATAGCTCTGGTGAACTGCATACACTGGT

ACAAGAGAAGAAA●TCTCTGCACTGAATCAAATACTCACAGCC●GGCTTTCTCCAGAA
TCTACCTTCTTTTAACAGTATTCACTGTTATAGCAGTGTCTACGCTATACACACACGTGT
TGGTAACTAGAAGAGTGGTAAAACTGATTAATTTCCATTTGCTTTTCAGCAATCATTTTA
GCATGTGGCTTGCTGCATGCCTTGGCCTTTATTATTTTCTTAAAATAGCTCATTTTCCTA
ACTCTATTTTTGTTTACTTAAAGATGAGAATTAACCAGGTGGTTTCAGGGACTTTGCTCA
TGTCTTTGGGCCTCTTGTTTCTAAACACTCTGCTGATAAACTCATACATTGATACCAAGA
TAGATGACTACAGAGAACATCTACTGTATGATTTCACTTCGAATAATACTGCTTCATTTT
ACAGGGTTATTTTAGTCATTAACAACTGTATTTTCACATCTATACCCTTTACACTTTCCC
AGTCCACTTTTCTCCTGCTCATCTTCTCCCTGTGGAGACATTACAAGAAGATGCAACAGC
ATGCACAAAGATGCAGAGATGTCCTTGCAGATGCCCACATCAGAGTCTTGCAAACCATGG
TCACCTATGTCCTACTCTGTGCCATTTTCTTTCTGTCTCTTTCCATGCAAATTTTGAGGA
GTGAGTTGTTGAAGAACATTCTTTACGTTAGGTTCTGCGAGATTGTTGCAGCAGTTTTTC
CTTCAGGACACTCCTGTGTCTTAATCTGTAGAGACACAAACCTGAGAGGGACCTTTCTTT
CTGTGCTATCGTGGCTGAAGCAGAGGTTTACATCATGGATTCCTAACATAAATTGCAGAT
CATCTTGCATATTCTAAAAGAAACTGAG

SEQ ID NO:162

Mouse T2R30 amino acid sequence

MTYETDTTLMLVAVGEALVGILGNAFIALVNFMGWMKNRKIASIDLILSSVAMSRICLQC
IILLDCIILVQYPDTYNRGKEMRTVDFFWTLTNHLSVWFATCLSIFYLFKIANFFHPLFL
WIKWRIDKLILRTLLACVIISLCFSLPVTENLSDDFRRCVKTKERINSTLRCKVNKAGHA
SVKVNLNLVMLFPFSVSLVSFLLLILSLWRHTRQIQLSVTGYKDPSTTAHVKAMKAVISF
LALFVVYCLAFLIATSSYFMPESELAVIWGELIALIYPSSHSFILILGSSKLKQASVRVL
CRVKTMLKGKKY

SEQ ID NO:163

Mouse T2R30 nucleotide sequence

AAAAATGTTCATTGTTTATCTAAAATTCAAATTTAACTGAGTGCCCTACATTTTTATTTA
TTCAATCTAGTAGCTGTACTGAGGTTATTAGTGTGATTTCTGAAGCCCAAATTTGTAAAA
CTTAGCCTCAGATAAACAGCTTGAGACCATGGAAAGTAATTTGGTAAA*TTTGCATCTTAG*

CAAATAGTAGCTC[redacted]CCTAAATTAACTGTGTGTAGAAAAGAATC[redacted]CTGCGGAGAAGATA
AATGGACATACAATATCCAGGCTAAGGATTGCCAAACACACTGTTTTTAAGACTAATTGA
GATTTAGATAAACTATCTACAGTCTTCATGTATAATTCTCATCTTCATCACAAGACAGAC
TTCAACTTAAGGAGGTAAAGACAAGGACAGCGAACCCTAAACAGCCAAGTGTAGAAACCA
AACTGCATCAAATCAGCCAGAAACTAATTGGATACTTCTCTACTTTAAA**ATGACATACGA**
AACAGATACTACCTTAATGCTTGTAGCTGTTGGTGAGGCCTTAGTAGGGATTTTAGGAAA
TGCATTCATTGCACTGGTAAACTTCATGGGCTGGATGAAGAATAGGAAGATTGCCTCTAT
TGATTTAATCCTCTCAAGTGTGGCCATGTCCAGAATTTGTCTACAGTGTATAATCCTATT
AGATTGTATTATATTGGTGCAGTATCCAGACACCTACAACAGAGGTAAAGAAATGAGGAC
CGTTGACTTCTTCTGGACACTTACCAACCATTTAAGTGTCTGGTTTGCCACCTGCCTCAG
CATTTTCTATTTATTCAAGATAGCAAACTTCTTCCACCCTCTTTTCCTCTGGATAAAGTG
GAGAATTGACAAGCTAATTCTCAGAACTCTACTGGCATGTGTGATTATCTCCCTGTGTTT
TAGCCTCCCAGTCACTGAAAATCTGAGTGATGATTTCAGACGTTGTGTTAAGACAAAGGA
GAGAATAAACTCTACTTTGAGATGCAAAGTAAATAAAGCTGGACATGCCTCTGTCAAGGT
AAATCTCAACTTGGTCATGCTGTTCCCCTTTTCTGTGTCTCTGGTCTCCTTTCTCCTCTT
GATCCTCTCCCTGTGGAGACACACCAGGCAGATACAACTCAGTGTAACAGGGTACAAAGA
TCCCAGCACAACAGCTCATGTGAAAGCCATGAAAGCAGTAATTTCCTTCCTGGCCCTGTT
TGTTGTCTACTGCCTAGCCTTTCTCATAGCCACCTCCAGCTACTTTATGCCAGAGAGTGA
ATTAGCTGTAATATGGGGTGAGCTGATAGCTCTAATCTATCCTTCAAGCCATTCATTTAT
CCTCATCCTGGGGAGTAGTAAACTAAAACAAGCATCTGTGAGGGTGCTTTGTAGAGTAAA
GACCATGTTAAAGGGAAAAAAATATTAGCATCATGAGCATATCTGAAGAAAAACTATCAC
TTTCTAAGAGAAAGGAAGACACGATCATTATCCGTCCTTTTCACATGAATATTGATTTCA
TGCAGTGACATCCTCTTAACAAACTTAAATTGAACCTTGAGAAATCTCATATACAGCAAC
TTTGCATGTCTCTATCTCTGCTTTTTCTCTCCTTTTCAATATGAGTTGACATAAAAAATA
ATTTTCAGAACAAATTATAACAGAAGAAAGGGCATTTTCATAATCAGTTCTGAATCACTC
CTCCAAATGCAAAGCTGCCTGACAAATTCAAAACAATTGTAACAGCATCTCACTGTCGTT
TGCATTCTTTGGAAAAGCAGGTGGTTTGTTCTTGGAGCCTGGCTTAGAGTTTTCTTCTTA
GACCATTGAATTATGTTCATGATTGGAGAAGAGTCAAGTACCAAGTAACAATTTTTATTG
TGAAGATGGGTGTTCATCATGTGATTTTGGCTGGCCTGGAACTTGTTATGTAGACTAGTC
TGTCATCAAACACACAAAGATCTGCCTGCCTCACCTGCCAGTTCTAGGATTCAAGGAATG
CACCACCACAGCTTGTTCAAGTGACAATTCTTACAAATGTTTTAGAAATAAATAATATAC
TAGAAATTAACACTGAATGTAAGTGCTGTTTAGGTATAAATTATGATTAAATGTTATAGT
TAGAAAATTATTTAAGATTATAGATCAGTGATGAAAATATTCTAGAATAAGTTTTATGAA
GAAACTTTTATAAAGAAACTGGAAAAAAATCTCTTGATTGCATATTGAAACAAATTTCTC

CAAAAAGAACAC[redacted]CAAATTTGCTCTAGACATCTAGACTGTAT[redacted]AACAGTGAATATGA
AAATATCATAACAGGATATAGCCTTTAGTATTGAAGACAGGTTCATCTATATTAAACCTG
CATACATACCTAAAAGACTAAGTCAATATCCCACAAACATATTTGCACTATCATGTCTAT
TGAAACACTATTCATAGTAGCTAAAATATGGCACAAAACTAGACATTCATCAATAGATGA
ATCAATAAAGCAAATGTACATACACAAGATGAAATTGTATTCAGGCATAAAGAAGAATGC
AGTCATGTCATTAGCAAAAACATAAACAGAATTGGAGGTCATTGTGATAATTGAAATAAA
CCAGACCTGGAAAAAACAAAACCTGTGTAATTTTTCTGAAGTAGAGAATATACTCTTGGA
TGGATAGATGGGTACTGTTATAGTATAAAATGTGTGTGTGTGTGTGTGTGTGTGTGTG
TATTTCATGAAAGCAAGAATGGGACTGCTTAGAGAAAGAAAAGGACAAACAGGTGAAGGG
GTGAAAGAAAAAGGCAATGACAAGGAGTAATGATATGAGCAAAGTACCATTATTAAACAT
GTGACAATATTATATAGAAACACATGATTTTGTGTGCCTACCAAAACTGGATAATAATTT
TTAAAATGTATCTATTAAAAGGAAAGAAAAGAAAGTGCAAGCCCAGGAAAGGGAGAAAAG
GAAACAATGAGAGAGAAATGGAAAATGGTGAGAAGTGAAGAGAACAAAAAGAAATGGAGT
AAGTGTGGCCAGGAATGAAGGATCTCAGCTATAGTTATCCCAGTACGGTAATACAAATCT
GTGACTCCAGCACTTGACAAGGCTGAGAGATGTGAGAGAGGGCCAGTTAACAACCAGTCT
GGGCTTATTCCAAGAGATAAGAAGATTGGGGGAAAGTATGTAGAAGGGTTTGGAGGGAAG
AGAGAGAAGAGGGAAATGATGTAATGATAGTACAAATCAAAAGTTATTTTTTCTAAAAAA
GCAATGGGACAGGAAACCAACCTAACAAGTAAAGGTGCTTGGTTCACAAGACCAGCAACC
TGAGTGCATCCTTGCTAGAATGAAATTGGCCTTACTCTGGAAAGCTTACTTCCTCAGTGT
ATTCATTGTTAAAATTCATGTGGAGATTTTAAAGAAAAAAGGAAAAAAAAAGTTAAATGG
TAGATTTGTGTAGGGGAATATTCCCCTAATTAATTGATTAGATAATAAAGATGACAAGCA
AATTGCTGTGCAAAAAGGAAGACAAGGTCTAAGAGGGGAAGAGGGGACACGGGAGGAAAA
AAAACGGCCCTTTTTAAAGCAAGGTGGGGAGTGAGGGAAGCGAGATGTAGACAGGGAACT
GTTAGACCTGGTGGCAGCTTCTGCCACCTGAAGATTTTCAACATAGTATAGTTCATGAGT
TTAGGAAGATATGTTCCCTGCCCAGCGGTTGTATCATCTGTTGATTTTAAACTAAGATTG
TCTGGTGTTTTCCATTTGCGGAGACTCAAGTAGACCAAAGGGAAAGAATGAATTC

SEQ ID NO:164

Mouse T2R31 amino acid sequence

MYMILVRAVFITGMLGNMFIGLANCSDWVKNQKITFINFIMVCLAASRISSVLMLFIDAT
IQELAPHFYYSYRLVKCSDIFWVITDQLSTWLATCLSIFYLFKVAHISHPLFLWLKWRLR
GVLVVFLVFSLFLLISYFLLLETLPIWGDIYVTLKNNLTLFSGTIKTTAFQKIIVFDIIY

LVPFLVSLASLLI●LSLVKHSRSLDLISTTSEDSRTKIHKKAM●LVSFLILFIIHIFF
MQLARWLLFLFPMSRPINFILTLNIFALTHSFILILGNSNLRQRAMRILQHLKSQLQELI
LSLHRFSSLY

SEQ ID NO:165

Mouse T2R31 nucleotide sequence

CTGCAGCTTTCTAGAAATCTCACCAGAATGTCTTTGTGCAGCTTTAATAGTTCCTGGTTA
TACCTTGTCACATTATAAGCTAAGACATCTTTGGTGCCACAATATACTCTCACTAATCAG
AGAGATTAGACAGAAAAATAAGTTTCTTAACAACTGTTTTAGATAGGGTCATGAAATGAC
ATAAAACACCAATGCTAAGGCAATCCATTATGTTTTCTCATGAGGAGCCCATATGTACAC
TTGAGTGTGTCTTATTATTTCCCTGAGTGATTTTGTAATTTTATTAAACACTTAACTGTG
ATTCATACTAGTTAGTTCTGAAATTCTTTTCTTCATCAAAGCCATTAATCCTGGGGTTTT
TTAAATGGAGAACCCCAAAACAAAGTGAAATGTTGTGTGTGGAGCAGGCTGTCTTCCCAC
ACACTACCATGAGATGCTCATTCTGTAATTGTTCCCCGGAATAGGAAATGCCCTGAATTC
AGGCACACAAGAGCTAGTCTGTGCACCATGTCTGGTTCTTGCATTAATACCCACTTTTGT
CACGAAGCTTCATTGATTCGCATCTTCAGAAGCTGGTATCATTATTAGTTTCTTTCCTCA
GGTGACTCTGGnCCAAAATATTAnGGCGCCCTTTAAAAAAGTAAAACTACAAAATTTCTT
TATAATTTTCTTTAAGTTTGTTATAATATAGCATGACCTACACACACACACACACACACA
CACACACACACACACACACAAGTATGCCTCTCCTTTCCTTCTAAAAATCTCACTTAAAGC
AATTGTTTAGCTGTCTTCGAAGTCTAGACTGCCACTGTCGTGCTTCTAGCCAAAACAAAT
GCAACACATAAAATGATAGAGCTCAAAACTTAGGAATCTATTTAACTGTGAAGATCACGC
AAGCAAACCTGAGAAACCTCTAGAAGGAAACCACAGCAAATCACTGGAGAGAAGGTGTTA
ATCTAGTAAGAATAGTTTTTATTTTGGGTATCCTTTTGTAGATTGGTTAGTTCATCCAAA
ATCCAACTTGTTAGTTCTTCATAAATTGTAAGTGTCTCCAACATCAAAGCACCACTTCTC
TCTTTTCCCCTGTATGAAGATGCTTTAAGTACAGAGTTACTCTTTTTCTGTACTGACAGT
AATTTAAAAAAAATTGTTCACTCATTCTTTTTTGGTGTTGTTATTCTGTGTTCCTCAATGT
TATCTTTTTTTTTTCAAAACTTTCTTTTATAAAAAGTCATACACATAGCAAATGCAGTGC
ATGTTTATGGAATCCATAACTAACTTATTGAGACTTCTCCTAGTACTTTCTTTGAACAGT
AACAAAGATATCTGCTTCTACAGAGTGCAGTGTTTCAGGTGAGGAGGAACATATTATACA
AATCAGTGAAAAAAAAATCTGATTCAAATTTGTATTTTAATATATTTGACTTTATCACTT
CAGATATTACATCAATGGGAATTTTGAAGGCACACAAGTGATGATGTGGGCATAGAGACT
GTCTGTACTAGAATTTAATATTTCTTTTAAATATCTTTAAATAAAAATATGATGCTGTAT

TCATAAACAGAT[redacted]TTATAGATTAAGTATGAGATTAAAGTTGG[redacted]AAACAAAAGACAAAA
ACCTAGGACTAAGAATTTCCTTAAGTATGTGTGAATATCAACCTAATGGAGGAAGTTTCC
AATCAAAGCTGAAATTACAGTAAAAAGGAGGAAGATAAATATGGAAAAGGATGATTTTCT
GTGGAAGTTTGTTTGAGAACTGATCCACGAGACAAATTGCTAG**AAG**TGTGGATTCCCTTT
TACTATTCAACTGCTTATAGGACTGGATCAA**ATGTATATGATACTGGTAAGAGCAGTATT**
TATAACTGGAATGCTGGGAAATATGTTCATTGGACTGGCAAACTGCTCTGACTGGGTCAA
GAACCAGAAAATCACCTTCATCAACTTCATCATGGTCTGTTTGGCAGCTTCCAGAATCAG
CTCTGTGCTGATGTTATTTATTGATGCAACCATACAAGAACTAGCGCCTCATTTCTATTA
TTCTTACCGTCTAGTAAAATGCTCTGATATATTCTGGGTTATAACTGATCAACTATCAAC
ATGGCTTGCCACCTGCCTGAGCATATTCTACTTATTCAAAGTAGCCCACATTTCCCATCC
CCTTTTCCTCTGGTTGAAGTGGAGATTGAGAGGTGTGCTTGTTGTTTTTCTTGTATTTTC
TTTGTTCTTATTGATTTCTTATTTTCTACTGCTTGAAACACTTCCTATTTGGGGAGATAT
TTATGTAACCCTTAAAAACAATCTGACCTTATTTTCAGGTACAATTAAGACCACTGCTTT
TCAAAAGATAATTGTTTTTGATATAATATATTTAGTCCCATTTCTTGTGTCCCTAGCATC
ATTGCTCCTTTTATTTTTGTCCTTGGTGAAACACTCCCGAAGCCTTGACCTGATTTCTAC
CACTTCTGAAGATTCCAGAACCAAGATTCATAAGAAGGCCATGAAAATGCTGGTGTCTTT
CCTCATTCTCTTTATAATTCACATTTTTTTCATGCAGTTAGCACGGTGGTTATTATTTTT
GTTTCCAATGAGCAGGCCAATTAATTTCATCTTAACATTAAATATCTTTGCCTTAACTCA
CTCATTTATTCTCATCCTGGGAAATAGCAATCTTCGACAGAGAGCAATGAGGATCCTGCA
ACATCTTAAAAGCCAGCTTCAAGAGCTGATCCTCTCCCTTCATAGATTCTCCAGTCTTTA
CTAGAGGAACAGCTTAACAGGGAGACTTGGAAGGTCACTGGCAAATTATTCTTCTTTGAT
TTCTTTTAAGTACTGCTGAACATATATGAACTGTCCCCAGAGCATAGTGCTATCTTATGA
GAAGGATATCATCTCACAGTCTGGTTATAAAACACAAACCAATCTTTTTATAATTTCTTT
ACAGCATTGCTAATAAAAGACTTGTAGTCTCAAATATTTTAAAGAGAATAATTAATTTTA
TAGGCAAAAGGTATGAAATTACAATTCACAGGGAAGGTTCATGACTCCTTAGATATTAAA
GTTAATTGTAAGCCACAATAGGCAGAAGATGAGCAAAATGTTGATAGGAGATAAATAAAA
TCTAAAGTTACGGAGAAAAAAAACATCAACTTGCCTTTTAGATTACTTTAAAGCTCTCTC
TCTCGCTCTCTCTCTCTGTATCTACTTACTTTATATATACAAATGTTTTGTCTGCATGTA
TTTCTTTGCACCATATAAATGTCTAAGTATCCAGAAnGTCAGCAGAGGGCATCAAATTCT
CTGGAAAGAGAGTTACAAATTGCTGTGGGTAACACTGGGTGCTGGGAACTAACCTGAGTC
CTCTGCCACAGCAACTGCTCTTCCCTGCTGAGTCATGTTTTAAGTCTCCACAACTTAAAC
TCATTGTTGATGTGGTCATTGCATAATGATGAATTTACATTCTAAGGTTTGTATCATAGG
TAGGAGGGCTGGTTTTAATCATATTCTAATGTTCTTATACAAACCCAGGTTTTGTAAGAG
ACTGTATTCTATCATGAGACTCTTTCCCCACACCGCCAATGTAACATTTTTATTAATTTT

GAGGGGAATTTT[redacted]ACAGTGTACCCTGATCACCCTTGCTTCCC[redacted]TCCTTGCAGGTCTAC
CCTCCCACCATTGCTCAATCCCCCCTAAAAGAGAGAGAAACAAACCATGTCCAATTTGTG
TTGGACACATACTCAGTGGAACATGGCCAAACCCCTAGTGAGCAGTTCCTTAAAGAAAAC
TAAGCTGCCTCCCCACCACTACCACCATAGGGCATTAACTGTGAAGAGCTACACTTTAGC
TATTTTATCACCAATTTAAAAGACTGTCTTCAATAGCTTCCTCTATGGACTGTTTCTGGT
TTTAGTGGGACAGGGAGAAGGGGTCAAGAGGTTGTCACAGAAACTTTTGATGTCTCTTAT
TCTCAGTTAAAGTCCACTGCAAAAGAAGTCTGCTGGCTCTAATAAAGCTTGCAACAGCAT
GGGCCAGTGACATCATCATGATTTCTGGCAACAATATGGACCACAAATATCATGGCTCAG
GTGGCATTACGGACCACAGACATCAACATGGTCTCTGGCAGCAAGAACCAGAATCTTTTG
AGGAGGCTTCATTCAGAAAATGAATTTTTCTTCATCCCAGATATACTGATGTTGCTCAAT
CAGAGTATTAGTATGGTTGGGCACCATATTTGGGGACAGGACCTTCAATATTTCCAGGCT
GCTGTGTAACACATTATCTTTAGTGTCAGGTGCCCTTAGTGTCAGGACATGACCATCATG
TATGCGCCTGTGGGCAGAAATACATCTTTGTACTTTCTTACACCTAGCAGGGTGAGTAGC
AGGAGCAGCGGCATTAATACTTCCATACCTCTGGGCAGCCTATCAGGTATCATCTAGGCA
AGGTAAGCCCAGTAGTGGCCCAAGGCTCCTGGTGTCTACTTGGCAACAACATGCTCCTTT
GTCTGCACTGCCATATCTATGGCTGGTTCTCCATCCCTAGTTCTGCTTCTCTCAGGTTTT
ATACGACTCTATTCCACATTCTATTTTTCCAGTTCCATGAAACCAGTGTTTAAAAGTATC
ATCCCATAAGACCGGCCTTTTAAAGGTTATTCTGGAGATATTGCAGAGTCTGCAG

SEQ ID NO:166

T2R Family Consensus Sequence 1

E(F/A)(I/V/L)(V/L)G(I/V)(L/V)GN(G/T)FI(V/A)LVNC(I/M)DW

SEQ ID NO:167

T2R Family Consensus Sequence 2

(D/G)(F/L)(I/L)L(T/I)(G/A/S)LAISRI(C/G/F)L

SEQ ID NO:168

T2R Family Consensus Sequence 3

NH(L/F)(S/T/N)(L/I/V)W(F/L)(A/T)T(C/S/N)L(S/N/G)(I/V)

SEQ ID NO:169

T2R Family Consensus Sequence 4

FY(F/C)LKIA(N/S)FS(H/N)(P/S)(L/I/V)FL(W/Y)LK

SEQ ID NO:170

T2R Family Consensus Sequence 5

LLI(I/F/V)SLW(K/R)H(S/T)(K/R)(Q/K)(M/I)(Q/K)

SEQ ID NO:171

T2R Family Consensus Sequence 6

HS(F/L)(I/V)LI(L/M)(G/S/T)N(P/S/N)KL(K/R)(Q/R)

This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☑ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☑ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☑ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☑ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.